Scuola Internazionale Superiore di Studi Avanzati - Trieste

DOCTORAL THESIS

The emergence of structure from continuous speech: Multiple cues and constraints for speech segmentation and its neural bases

Author:
Ana FLÓ

Supervisor:
Jacques MEHLER

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

DOCTORAL THESIS

The emergence of structure from continuous speech: Multiple cues and constraints for speech segmentation and its neural bases

Author:
Ana FLÓ

Supervisor:
Jacques MEHLER

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in

Cognitive Neuroscience

September, 2016

SISSA

Abstract

Cognitive Neuroscience

Doctor of Philosophy

**The emergence of structure from continuous speech:
Multiple cues and constraints for speech segmentation and its neural bases**

by Ana FLÓ

This thesis studies learning mechanisms and cognitive biases present from birth involved in language acquisition, in particular in speech segmentation and the extraction of linguistic regularities. Due to the sequential nature of speech, uncovering language structure is closely related with how infants segment speech. We investigated infant abilities to track distributional properties on the stimuli, and the role of prosodic cues and of memory constraints. In two experiments we investigated neonates' capacities to segment and extract words from continuous speech by using fNIRS. Experiment 1 demonstrates that neonates can segment and extract words from continuous speech based on distributional cues alone; whereas Experiment 2 shows that newborns can extract words when they are marked only by prosodic contours. Additionally we implemented a method for the study of the dynamics of the functional connectivity of the neonatal brain during speech segmentation tasks. We identified stable and reproducible functional networks with small-world properties that were task independent. Moreover, we observed periods of high global and low global connectivity, which remarkably, were task dependent, with stronger values when neonates listen to speech with structure. In another set of experiments we studied memory constraints on the encoding of six-syllabic words in newborns using fNIRS. Experiment 4 demonstrates that the edge syllables of a sequence are better encoded, and Experiment 5 goes beyond by showing that a subtle pause enhances the encoding of intermediate syllables, which evidences the role of prosodic cues in speech processing. A final group of experiments explore how information is encoded when it is presented continuously across different modalities; specifically if an abstract encoding of the sequences' constituents is generated. Experiments 6-9 suggest that adults form an abstract representation of words based on the position of the syllables, but only in the speech modality. In Experiments 10 and 11 we used pupillometry to test the same in 5-month-old infants. Nevertheless results were not conclusive, we did not find evidence of an abstract encoding.

Acknowledgements

This thesis is the direct product of a bunch of coincidences. For some reason I decided to study chemistry, even knowing I wanted to investigate the brain. By the end of my degree I was confused about everything: high cognitive functions or motor control? the human brain or a leech? leeches or coordination complexes of Ruthenium? research or a life in Patagonia? One day, going to a lesson, I saw a poster announcing a seminar by a certain Prof. Jacques Mehler, something about language and infants. However, when I looked more carefully I noticed that the lecture of Prof. Mehler had been the week before... but something intrigued me and I wrote down the name. Jacques Mehler had studied chemistry and (even if he was Jacques and not Santiago or Jaime) in my same university, and if that was not enough, he used to work in Trieste, a beautiful place where I had spent a couple of years during my childhood. I asked around more about Prof. Mehler and, suddenly, all my enthusiasm fell off: he was retiring. But I however came to SISSA, and it was here that I discovered he had won a grant and he was staying for another five years. I perfectly remember my first meeting with Jacques and Marina. They made me feel very relaxed, they listened to me, and they stimulated my curiosity. At the same time they offered me to do a PhD with them, they gave me a long list of books and papers to beat my ignorance. I am very grateful to Jacques and Marina for this opportunity. Thanks Jacques for pushing me into those long days of reading without exactly knowing what I was looking for (I don't think I've understood it, but somehow the present work came out and this is because of your subtle guide), and thanks for conveying your interest on the research with neonates! The way has been long and not always easy, but I've learned from you. Marina, thanks for convincing me about the importance of prosody (I think sooner or later the ITL project will work!). Thanks to both for the nice dinners and good wines.

Special thanks to Alissa Ferry. With her, we planned and discussed many of the studies in this thesis. I wouldn't have been able to do this work without her support, her advice, and attentive corrections (Alissa, I promise to keep working on my writing and in suppressing auto-destructive behaviours!). We shared alone the early morning trips to Udine during more than two years, till the arrival of the French saviour. Thanks Perrine Brusini for the company during the Udine missions, but specially for your constant and genuine preoccupation, for the (very!) honest comments, and for filling my linguistic ignorance. And girls, I'm happy to have shared more than science with you.

When I talk about the testing mornings in Udine it is impossible to forget Prof. Macagno, I am grateful for his help, and I admire his dedication during all these years.

I want to thank all the people that were part of the lab during my stay in SISSA. For their comments and support, and for all the problems and charm derived from

working together. Maybe, we could all be defined as peculiar characters, but (and maybe because of this) the atmosphere of the group has always been delicious, I felt I could trust any of you. In particular thanks Julia for tolerating my sometimes autistic mood while working and for being always predisposed to help; thanks Hanna for your acute comments and criticisms; thanks Amanda for transmitting a devastating energy and optimism; thanks Alan for your support and for trusting me. I want to specially thank Milad and Michela for all the discussions on miscellaneous topics that during working hours were camouflaged under coffee breaks, and for the many crazy moments. Un gracias especial a Yamil. We started together this adventure blind to the future, and we are arriving to the end. I could try to explain what it was to go through these years together in many ways, but I think you defined it in the best manner a few days ago when you told me: *"If I'm with you and for example I have the mouth full of food I am relaxed because I know you can answer in my place to whatever question"*. I feel the same. Gracias, for the support in many aspects, I'm still trying to learn from your approach to life and its complications. *Imitemos el ejemplo de este varón argentino, y siguiendo su camino gritemos de corazón: ¡Viva Yamil! ¡Viva Yamil!... ¡Yamil, Yamil que grande sos!*.

The lab would not have held without the work of Marijana and Francesca. You made things work! with a continuous work, solving problems even before they appeared, tolerating all of us with our whims, and above all always with a smile. I also want to thank Alessio, in silence but always there, ready to solve any technical problem that appeared.

Besides people and time, to do science now a days there's another crucial ingredient, money. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/European research Council Grant Agreement 269502 (PASCAL) (to JM).

In general thanks to SISSA for all the scientific and non scientific meetings, which make it a great place to do research, from a scientific and human perspective.

But I would not have survived to the PhD if I would not had found amazing people in the way. Even if many of you have already left Trieste, and now it is my turn to leave, many friendships and great memories remain. It is the most precious thing I bear with me after these years. I had promised to myself not to do long lists of names but now I'm in troubles! this could become too long and emotive, thus I'll do what I didn't want to do... los parceros, Duvanchito y Fra, salsa, mar and... dov'è Tardizzi? Iga, I couldn't have been more lucky for going to your office the first day I came to SISSA, Iga and Angus you are synonymous of optimism and adventure. Waz, what to say, (una enseñansa memorable: comer el kiwi con cascara esta bueno porque comes y te lavas los dientes al mismo tiempo, lo dice todo ¿no?) te extraño! Alex, Victor han sido mi apollo en las malas y compañeros en las buenas. La famiglia patata, Alessandro, Maria, Alejandro. Thanks

to all the good friends I made in SISSA e a tutti i Triestini (e Triestini per adozione) che ho conosciuto in particolare durante gli ultimi due anni. Un grazie speciale (in modalità post-acknowledgement?) per Isa, che mi ha fatto scoprire quanto buone sono le burrate!

With time I realized I received a very good education during the university years in Argentina (always surrounded by that characteristic *quilombo*, which is actually part of the training). I never paid a cent for it, as all other students. But for this I should thank so many people and so distant from me that becomes an empty gratitude. I hope to be able to say thank, returning home in the future.

Concluyo con un gracias a mis viejos por haber siempre estado presentes y al mismo tiempo por haberme dejado espacio cuando me servía, por aceptar mis largas ausencias sin reclamos, y por la paciencia a la eterna duda. Ahora cuento los días para abrazarlos.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
BOLD	Blood Oxygen Level Dependent
C	Mean Cluster Coefficient
Ck	Betweenness Centrality
E	Global Efficiency
EEG	Electroencephalogram
fNIRS	functional Near Infrared Spectroscopy
Hb	Deoxyhemoglobin
HbO2	Oxyhemoglobin
HRF	Hemodynamic Response Function
IFG	Inferior Frontal Gyrus
MEG	Magnetoencephalography
MTL	Medial Temporal Lobe
NIRS	Near Infrared Spectroscopy
PCA	Principal Components Analysis
ROI	Region of Interest
sw	Small World
STG	Superior Temporal Gyrus

Chapter 1

Introduction

Human language is one the most salient cognitive abilities that distinguish humans from other species. It is much more sophisticated and flexible that the communicational system of any other specie. It allows the transmission of variety of experiences and thoughts, making possible cultural transmission. These are probably some of the reasons why human language has received so much attention since the beginning of philosophy.

The *nature vs. nurture* argument around language can be tracked back in time. The Egyptian Pharaoh Psamtik I sought to discover which is the natural language of humans, and to do so he gave two newborns to a shepherd with instructions of not talking to them. When the infants were presented to him two years later one of them said something like “bekos", which in Phrygian means bread. Psamtik concluded that there is a natural language of humans and that it is Phrygian. The existence of innate ideas was sustained also by Plato, and still by Descartes centuries later, but that it was strongly refuted by Locke, setting up the *innate ideas vs. tabula rasa* debate in philosophy

> *Let us then suppose the mind to be, as we say, white paper void of all characters, without any ideas. How comes it to be furnished? Whence comes it by that vast store which the busy and boundless fancy of man has painted on it with an almost endless variety? Whence has it all the materials of reason and knowledge? To this I answer, in one word, from experience.*

After centuries of years the question regarding which aspects of language are innate and which ones are acquired still does not have an answer. Infants are faced with this continuous flow of sounds that is speech, and after only a couple of years they start producing utterances in the language they have been immersed. They are able to combine the constituents of language in new ways in order to communicate their own needs or ideas, and remarkably, language learning in infants occurs fast and naturally, without explicit instruction.

In this thesis we assumed that there are specific biological constrains for language acquisition. The work is focused on one of the first problems infants face during language acquisition: segmenting the continuous flow of phonemes that is speech. Infants need to identify its constituents —syllables, words, sentences—

and moreover to extract the structure underlying it. The aim of this thesis is to dissociate specific language learning mechanisms and “innate knowledges", from skills derived from infants’ interaction with the environment. For the sake of answering this question the majority of the studies in this thesis are conducted with neonates, the population with less linguistic experience that we have access to. Moreover, we used neuroimaging to explore the neural bases of this learning mechanism.

Before going to the individual experiments I will make a general introduction and motivation for the studies presented in this thesis.

1.1 The language faculty and its acquisition

There are still two fundamental open questions —and points of controversy— around human language that I would like to remark. First, what (if anything) makes human language unique? and in more general terms, do humans have cognitive abilities that rely on qualitatively different mechanism than other animals? And second, how do infants acquire their language? Is everything learned or do “innate knowledges" exist?

Skinner, as behavioural psychology, thought that infants learn their language as consequence of the interaction with the environment by positive reinforcement (Skinner, 2003), but this idea was strongly criticized by Chomsky because being incompatible with language’s features and its flexibility.

The development of generative grammar theory in linguistic by Chomsky (Chomsky, 1957) meant a formalization of the study of language’s structure, which evidenced the existence of specific rules determining grammar —even though knowing a language does not imply an explicit knowledge of the rules governing it. Importantly this formalization triggered a fundamental observation and subsequent question: the input infants receive is not enough to generate such linguistic knowledge; then, how do infants master language in such a short time? Chomsky’s answer was the existence of some innate knowledge, and proposed the existence of an Universal Grammar (UG). The basic idea under the UG hypothesis is that all languages share some universal principles that are core knowledge, whereas the role of the input limits to a set "on" or "off" a number of finite parameters (*Principles & Parameters theory*). The content of this grammatical core knowledge has changed along years, till latest versions proposing a few operations that would make possible recursion in language (*Minimalist Program*, Chomsky, 1995). Besides the specific aspect and changes in the theory, what results crucial for us is that it poses the existence of something like a "language organ" in the brain, meaning that language is settled on specific mechanisms. Moreover, it implies that the human mind works based on symbols and abstracts rules, in a qualitatively different way than other animals cognitive systems.

Chomsky's theory challenged Skinner's ideas and meant a new perspective that lead to look for the biological bases of language; nevertheless, the biological foundations of the theory were not convincing.

The stronger theoretical arguments against UG is its evolutionary origin. Biologists have refuted the possibility of something like a "language organ" and the disconnection it poses between human cognitive abilities and other species, from the initial state of the theory (Lieberman, 1984; Bickerton, 1992). In newer formulations Chomsky and allied scholars have proposed that language appears as consequence of one or more mutations or gradual natural selection that enabled to process the hierarchical and recursive structure of language (Hauser, Chomsky, and Fitch, 2002). Notwithstanding their attempt, this reduction in what characterizes human language is far from solving the problem.

Another criticism to UG regards its implementation in the brain and comes from a connectionism approach. Connectionists believe that learning is a consequence of changes in the connections due to the input, meaning there are not specific learning mechanism. Language (as any other knowledge) can be acquired by innate but shared across domains learning mechanism. Under this perspective the amazing human learning capacity is the result of a higher computational power, and not of qualitatively different mechanism (e.g. Joanisse and McClelland, 2015). Something should be notice: the main aim of connectionism is to model networks that could support learning, in other words to look how learning is implemented in the brain. Connectionism does not intrinsically go against UG (see Fodor and Pylyshyn, 1988 for a discussion about symbolic learning vs. connectionism).

In sum, UG does not seem suitable from an evolutionary perspective, and there is no theory supporting how it could be implemented in the brain. In the attempt to answer the question of language origin, other approaches have been taken. For example Christiansen and Charter (Christiansen and Chater, 2008) brought to the extreme the idea of language as consequence of learning and memory constraints. They proposed that evolution cannot explain the origin of UG (or innate language capacities) and that instead "language is shaped by the brain". Under their view, language structure is consequence of non-linguistic constraints amplified by cultural evolution: pre-existing neural mechanisms were recycled, and language emerged restricted by sensory-motor constraints and cognitive limitations in memory, learning and processing. This approach represents a growing perspective trying to fill the gap between the emergence of a system with the complexity of language, and some operational principles of the brain that cannot be radically different from other species. Nevertheless —as it has been remarked in numerous comments to their paper— they mislead and neglected some important factors. There are two main criticisms to their approach. First, that biological evolution does enable adaptation to a changing target as language as long as there are common regularities (Barrett, Frankenhuis, and Wilke, 2008). Thus, adaptation can occur toward abstract features, giving place for example to useful

learning biases and constrains for language acquisition (Fitch, 2008). Second, that the main goal of language is communication, hence the driving force is not only learnability, but evolution puts pressure for an appropriate cognitive infrastructure for communication (Ruiter and Levinson, 2008).

This overview tries to show how, despite a big effort for understanding the origin and biological foundations of language has been made, not so much progress has been done. Some of the difficulties to answer these questions rely on the following facts (see Hauser et al., 2014): 1. Language is subject to cultural evolution, meaning first, that it is not easy to determine which aspects derived from cultural evolution and which ones from biological evolution; and second, that because language changes fast, it is difficult to determine what represents an evolutionary advantage. 2. Related with the former, there is no agreement about the forces that may have pushed for its evolution. 3. It is not possible to have fossil data of ours ancestors language's capacities. 4. Language is unique to humans, thus there are not appropriate animal model —even if comparative studies can be done and are actually very informative. 5. The link between language abilities and genes is still far to be known (even if some genes associated to language abilities have been identified).

To wrap up, we can say that nowadays there is general agreement that humans have some specific capacities for language acquisition, but the main questions remain still open. How has language evolved? and, to which extent does it rely on specific mechanism? Does it arise from pre-existing cognitive capacities or there was a qualitative change in the human mind compared with other species? The topic is intricate, but what it is evident is that it requires an interdisciplinary approach, that joins knowledge from linguistic to biology passing trough computer science and anthropology. We will address the problem from a developmental and cognitive science perspective. We will try to identify some general and specific language learning mechanisms during the very first stages of language acquisition in infants and we will investigate its neural basis.

1.2 Language acquisition in pre-verbal infants

The study of human development offers many insights to the language problem. It is possible to study what infants know and what they have to learn, what they learn independently of the cultural environment in which they are raised and what depends of the input they receive, and moreover what they can or cannot learn at each stage during development and the changes in learning strategies that occur.

In the language domain infants show amazing capacities that vanish at some point during childhood —performance in a language acquired during adulthood never riches the level of the mother language. This denotes the existence of the so called *critical period* for language acquisition. The critical period is the outcome

of a neural system ready for language learning in a specific moment during development that later in life becomes less fit to this task due to maturation and to the linguistic input received (Newport, 2002). The critical period shows clear evidence of the strong biological bases for language acquisition, but leaves open what is exclusively consequence of biological constrains and what depends of the input.

Acquiring a language implies acquiring all its components: its phonemes (phonetics), the words and the matching with their referent (semantic), and how the units are organized (morphology and syntax). The time line of this learning process is well delimited with all infants acquiring specific aspects of language at the same age, and numerous studies have revealed that infants are particularly sensitive to linguistic stimuli since very early in life.

For example, infants at birth are able to distinguish languages belonging to different rhythmic classes (Mehler et al., 1988; Nazzi, Bertoncini, and Mehler, 1998), and prefer to listen to their mother voice than other voices (DeCasper and Fifer, 1980; Querleu et al., 1984), showing that infants process linguistic stimuli even while being in the womb.

Other researches suggest that speech is treated in a distinctive way than other auditory stimuli since very early. In series of experiments Benavides-Varela et al. showed that neonates can form a short term memory of a bisyllabic word (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012), and that linguistic stimuli but not music, interfere with this memory (Benavides-Varela and Mehler, 2012), pointing out that the bases of a memory system for word learning is present since birth. Another study shows that verbal labelling, but not tones, enhances categorization at 3-month-old (Ferry, Hespos, and Waxman, 2010), showing that language occupies a special role that extends to cognition in more general terms.

Nevertheless this high sensitivity for speech processing and the salient role of language since birth, infants do learn a lot from the input they receive. In another experiment Ferry and colleagues found that while categorization is promoted by human vocalizations at 3 and 4-month-old, it is promoted by lemur vocalizations only at 3-month-old (Ferry, Hespos, and Waxman, 2013), demonstrating a further specialization in human vocalization around 4 months. More evidence of the fundamental role of the input, but at the same time of the existence of innate bases for speech processing comes from the acquisition of phonemes. Since birth, phonemes perception seems to be categorical (Eimas et al., 1971; Eimas and Miller, 1980; Dehaene-Lambertz and Pena, 2001) denoting some type of "innate knowledge" or bias in perception; however, neonates discriminate all possible phonemes contrasts and it is only during the first year of life that infants become tuned to the phonemes of their own language (e.g. Polka and Werker, 1994; Kuhl et al., 1992; Werker and Tees, 2002; Werker et al., 1981; Maye, Werker, and Gerken, 2002).

The literature evidencing how infants learn from the input is vast and extends to various features of language. Infants learn regularities on segmental prosody, as demonstrated by a study showing that 4-month-old infants can already discriminate the stress pattern of their own language (Friederici, Friedrich, and Christophe, 2007); and moreover, infants extract the structure of the language they have heard: at 8 months, infants parse artificial speech with frequent syllables occupying the position that frequent syllables (as articles, determiners, pronouns) have in their own language (Gervain et al., 2008a).

But, what kind of learning mechanisms sustain this astonish infant language learning abilities?

1.3 Learning mechanism for language acquisition

The literature has distinguished three main learning mechanisms: statistical learning, rule learning, and more recently the use of perceptual biases.

Statistical learning, in the most general term, means tracking occurrences in the input and based on it to predict the outcome. Regardless of the poverty of the stimulus argument, speech is a rich signal; thus, potentially, regularities can be extracted based on the occurrence of the components. Research on statistical learning had a boom after the work of Saffran, Aslin and Newport (Saffran, Aslin, and Newport, 1996), where they found that 8-month-olds infants can extract words from continuous synthesized speech based only on the distribution of the syllables in the stream. In this specific case the units whose occurrence was tracked were syllables in speech, but statistical learning applies to any feature of language, and moreover to any feature of any stimulus in any modality. For example infants could acquire the phonemes or the stress pattern of their language by means of statistical learning; and infants learning from items' distribution has been extended to the non linguistic domain in the auditory modality (Saffran et al., 1999; Kudo et al., 2011) and to the visual modality as well (Bulf, Johnson, and Valenza, 2011; Kirkham et al., 2002). Moreover, even non-human animals are sensitive to distributional properties of speech (Hauser, Newport, and Aslin, 2001; Toro and Trobalón, 2005), clearly denoting that learning from the co-occurrence of the stimuli is a very general learning mechanism across domains and species. Something should be remarked regarding the domain generality of statistical learning. Statistical learning was born from a theoretical perspective opposite to supporters of the existence of domain specific mechanisms. However, there are not neural basis to believe that because distributional regularities are tracked across modalities, statistical learning works in an equal manner for all type of stimuli (Frost et al., 2015). The processing could occur in different cortical areas, with specific constraints in each.

While statistical learning means to generate associations between specific items and items are defined by superficial features of the stimulus, ***rule learning*** refers

to the capacity of generating a more abstract representation of the regularities. The fundamental aspect of rule learning is that it enables the generalization of a pattern to new instances. The capacity of generating abstract representations and a symbolic thinking has been considered a distinctive aspect of human cognition and it is a requisite for UG: language is learned by the acquisition of a set of rules. The first evidence of infants' capacity to generalize a pattern in language's structure was observed with 7-month-old infants. Infants learned ABA and ABB patterns and generalized them to new instances (G. F. Marcus, S. Vijayan, S. Bandi Rao, 1999). A more recent work shows similar abilities in neonates (Gervain et al., 2008b); and other studies suggest that this type of generalizations is not restricted to language, by showing that infants perform the same type of generalizations on the visual domain (Saffran et al., 2007; Johnson et al., 2009).

The traditional interpretation of the artificial grammar learning experiments mentioned before is that infants generate a rule from the input that later apply to new instances. This view has been more recently argued by some researchers that proposed that infants could have generated not any abstract representation but instead they may be sensitive to a salient property that was present in the stimuli, that is to say repetitions (Endress, Scholl, and Mehler, 2005). In more general terms this new perspective poses that learning could occur by the existence of ***perceptual biases*** (Endress, Nespor, and Mehler, 2009). Under this view perceptual biases are the outcome of a neural system that does not process the input in an equal way in all its dimensions. Some features may by more salient as consequence of how sensory information is integrated, or can be better encoded due to memory constraints. Endress and colleagues described two perpetual biases —or perceptual primitives— for artificial grammar learning. On one hand the detections of repetitions we talked before. On the other hand an encoding of the position in the sequences relative to the edges, with a better encoding of items occupying the first and last position.

But perceptual biases go beyond artificial grammar learning. Perceptual bias could for example explain universals in prosody. A study shows that adults are able to recognize prosodic units of unfamiliar languages (Endress and Hauser, 2010), suggesting that the perception of prosodic units does not depend of language experience. Moreover, the existence of a grouping bias, known as Iambic Trochaic Law (ITL), was described initially for music and afterwards extended to the linguistic (Hayes, 1985; Hay and Diehl, 2007; Nespor et al., 2008) and even to the visual domain (Peña, Bion, and Nespor, 2011). The ITL states that elements contrasting in pitch are grouped with the high pitch element at the beginning, while elements contrasting in duration are grouped with the longer element at the end. Perceptual biases bring an innovative perspective: language learning can rely also in constraints imposed by how the sensory and memory systems are built and how they develop. Notice that these constraints are not necessarily modality nor specie specific.

The three learning mechanisms described above have been traditionally considered independently because of theoretical motivations. In the last years intermediate positions have emerged. Generalization to new instances could be consequence of tracking regularities in the stimuli (statistical learning) through a neural system with intrinsic cognitive biases. Under this view, generalization would occur when the new items are consistent in terms of some salient features (Aslin and Newport, 2012). Moreover, differences in learning based on distributional regularities may differ across modalities (Frost et al., 2015), which may lead to some type of specificity in processing, even though through a mechanism, from a theoretical perspective general, as statistical learning.

Cumulative studies demonstrate that development is a conjunction between genetic or biological limitations and the interaction with the environment. This interaction is intricate, but its uncover will lay the foundations for the understanding of language acquisition. This thesis will try to provide evidence that language learning relies on a combined use of the three mechanisms described before, and that they should not be considered dissociated. Furthemore, learning from distributional regularities (statistical learning) can be susceptible to perceptual biases and to modality-specific limitations, that are consequence of brain's structure.

In this thesis we work with very young infants in order to identify cognitive biases in perception and memory constraints existing since birth. In particular we focus on the segmentation of speech that constitutes one of the first problems infants face during language acquisition. We investigated neonates abilities to segment continuous speech based on distributional (statistical learning) and prosodic (perceptual biases) cues. Moreover, we investigated how sequential information is encoded, and if it is affected by the presence of prosodic cues. We believe that infants capacity for extracting structure from speech depends on their ability to compute the distribution of different features across the stimulus, and also on the existence of innate perceptual biases, and memory constraints and limitations. As an example, we can imagine infants hearing a continuous flow with the structure *...AAxxxCBBxxxDAAxxxCBBxxxD...*, where *AA* predicts *C* and *BB* predicts *D*. If long distance dependencies are hard to track and memories of long sequences are not easily generated, it is implausible that infants will succeed at extracting this pattern, but they may extract the shorter sequences *CBB* and *DAA*. If instead prosodic contours delimit the *AAxxxC* and *BBxxxD* structures infants may succeed, because now what they need to track is that some sequences begin with *AA* and finish with *C*, while others begin with BB and finish with *D*. Experiments in this thesis try to shed some light in this line.

1.4 Neural bases of language

There is much evidence showing that language relies on specific neural circuits in the adult's brain (see Friederici, Oberecker, and Brauer, 2012; Friederici and Gierhan, 2013; Price, 2010 for reviews). This system is located over the frontal temporal cortex with a dominance of the left hemisphere. It includes the two well known Broca and Wernicke language areas and extends on the superior and middle temporal gyrus. Broca's area, located in the inferior frontal gyrus, has been traditionally associated with speech production; whereas Wernicke's area, located in the posterior part of the superior temporal gyrus, with the semantic aspects of language. A series of pathways connecting the different areas and the motor cortex are known (see Figure 1.1).

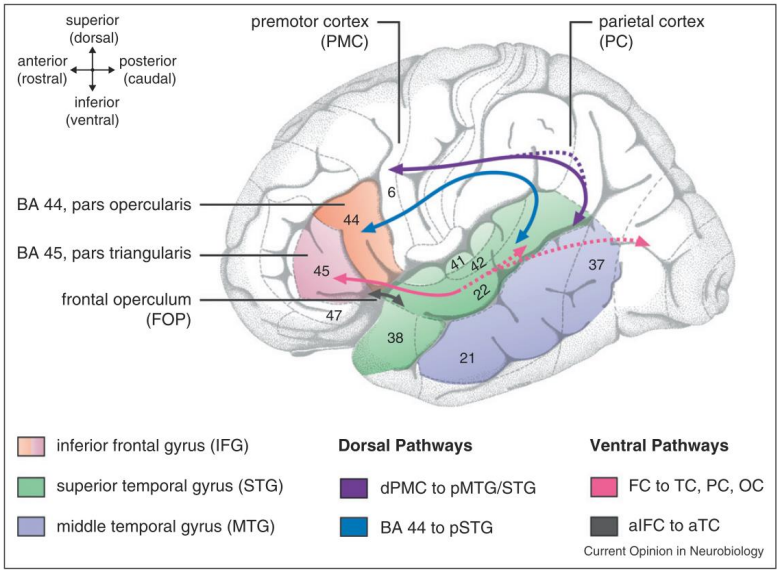


FIGURE 1.1: Language network. Adapted from Friederici and Gierhan, 2013.

Although there is no doubt that language processing and production is done by very specific circuits in the adults brain, many questions still do not have an answer. Which are exactly these structures and how are they organized? What does each of them compute? Answering these questions and uncovering the differences with other species will lead to the identification of distinctive properties of humans language and cognition. For example activation in Broca's area is certainly associated with the processing of complex sequential information, but many authors have suggested it responds to hierarchical structures (Bahlmann, Schubotz, and Friederici, 2008; Floel et al., 2009), and grammar (Musso et al., 2003; Friederici et al., 2006). If Broca's area is probed to process hierarchical information and a similar processing mechanism is not observed in other animals, this will give support to the existence of computational principles belonging to

the human brain.

However, there are questions we cannot answer by studying language processing and production in the adult brain. Are the neural circuits for language learning settled from birth? or does this organization consolidate during development as result of the input infants receive? To answer these, the neural bases of language have to be studied during development.

Neural development has traditionally been studied in humans by post-mortem studies and using animal models. Neurons differentiate and form a six layers structure like in the adults' neocortex already during the fetal period, but the connections are established much later (see Stiles and Jernigan, 2010 for a review of brain development). Some important connections are set before birth, but a big increase in the number of connections, synaptogenesis, and a subsequent selection of the relevant one, synapses pruning, continue during the first years of life. Furthermore, the myelination extends till adulthood. Moreover, these processes do not proceed in a uniform way. Synapses pruning occurs first in primary sensory and motor cortices, then in associative areas on temporal and parietal cortices, and later in higher associative areas over frontal cortices. This time course of the maturation, lead to the traditional view that establishes that the development of frontal areas as well as new functional systems as language, occurs as consequence of the interaction with the environment, and therefore that higher areas in young infants are not very active.

In the last years the development of imaging techniques has enabled great progress, questioning the classical view that development occurs from bottom to higher areas. This idea has been challenged in distinctive ways.

First, by recent studies showing that even at birth neural networks are not limited to local areas but include associative areas as well. One study with preterm infants described functional networks with a similar organization of adults: visual and auditory networks appear stable even in preterms, whereas somato-sensory, motor, default mode, and executive control networks emerge in the third trimester (Doria et al., 2010). Moreover, habituation studies have also shown an active role of frontal areas (Nakano et al., 2009; Mahmoudzadeh et al., 2013), and a recent study demonstrated the existence of top down regulation in 5-7-month-old infants (Emberson, Richards, and Aslin, 2015).

Second, specialization for the processing of some particular stimuli seems to be present from birth, and crucial for us, a left lateralization for speech processing was observed in 3-month-old infants (Dehaene-Lambertz, Dehaene, and Hertz-Pannier, 2002) and neonates (Perani et al., 2011; Peña et al., 2003) (see Dehaene-Lambertz, Hertz-Pannier, and Dubois, 2006 for a review).

Third, neuroimaging techniques have also made possible to describe anatomical changes during development. Importantly, numerous left right asymmetries in maturation have been reported, which may explain the advantage of the left hemisphere for speech processing (see Dehaene-Lambertz and Spelke, 2015 for

a general review). Asymmetries in the maturation of white matter tracks and in the development of the sulci have been described (see Figure 1.2). The arcuate fasciculus, connecting temporal and inferior parietal cortex with the frontal lobe (Broca and Wernicke areas) is further developed in the left than in the right hemisphere during the first trimester of life (Dubois et al., 2009; Liu et al., 2010, see Dubois et al., 2014 for review of white matter maturation). In a recent work Leroy et al. (Leroy et al., 2011) report a faster maturation of the right relative to the left superior temporal suculus, and moreover, this asymmetry is correlated with a faster maturation of the left inferior frontal region. Both areas are central elements of the phonological loop for auditory working memory in adults (Baddeley and Hitch, 1974), suggesting the maturation of this circuit may initiate very early in life. These anatomical differences may explain hemispheric specialization for speech processing very early in life.

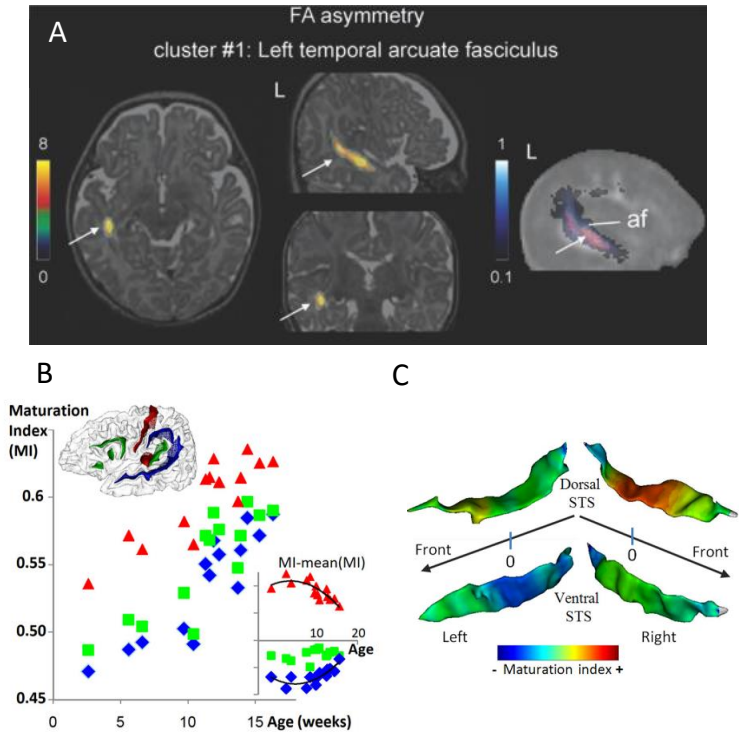


FIGURE 1.2: Asymmetries in brain development. A) Arcuate fasciculus left asymmetry. Adapted from Dubois et al., 2009. B) Maturation indexes for different areas. Primary cortices mature first (red), follow by the inferior frontal sulci and the planum temporale (in green), and finally the superior temporal sulci and the supramarginal gyrus (in blue). Adapted from Leroy et al., 2011. C) Right asymmetry in the maturation of the superior temporal suculus. Adapted from Leroy et al., 2011.

Brain development is the result of a complex interaction between, genetic, epigenetic and environmental factors, and infant cognitive abilities are the outcome

of these processes. Neuro developmental studies aim to identify the anatomical and functional differences with the adult brain, to track changes along development, and likely to understand the importance of environmental factors in each stage during development and to link the anatomical and functional architecture with cognitive capacities.

In this thesis we use functional Near Infrared Spectroscopy (fNIRS), a technique particular suitable for infants studies, to measure cortical activity in neonates during different speech processing experiments. We have two general goals. First, to study newborns' cognitive abilities, something that is otherwise difficult because behavioural measures are hard to get. And second, to identify the areas engaged in processing the information. Little work has been done to explore the neural bases of language learning mechanism as statistical learning and the perception of prosody. This thesis investigates the functional architecture of the neonatal brain while processing linguistic stimuli. We expect these results to help in the understanding of the link between the anatomical structure and maturation of the brain, and infants' cognitive abilities.

1.5 Aim of the thesis

This thesis is focused on those learning mechanisms and cognitive biases available from birth, that enable infants to extract the structure underlying language. Due to the sequential nature of speech, uncovering language structure is closely related with how infants segment or chunk speech. We believe that language acquisition is the outcome of the use of multiple cues. It is the interaction of what are generally called general learning mechanisms, meaning tracking distributional properties on the stimuli, and modality specific constraints in processing and memory.

In a series of fNIRS studies with neonates we explored the available mechanisms for speech segmentation and memory constraints operating since birth. Afterwards, we performed some behavioural studies with 5-month-old infants and adults to investigate the possible implication of memory constraints in the extraction of words from continuous speech and rule learning. In parallel, we used fNIRS recordings to examine the functional architecture of neonate's brain.

The thesis is formed by four experimental chapters plus a final chapter with a summary of the findings and a general discussion. The experimental chapters start with a theoretical introduction and revision of the bibliography and are articulated in the following way:

Chapter 2 investigates neonate capacities to use distributional cues and prosodic information to segment and extract words from continuous speech. We tested the use of the two mechanisms independently in two distinctive experiments using fNIRS. Experiment 1 investigates infants statistical learning

abilities, whereas Experiment 2 their ability to recognize words marked by prosodic contours.

Chapter 3 explores the functional architecture of the neonatal brain. We developed a functional connectivity analysis for the study of functional networks dynamics. We applied the analysis on the data of Experiments 1 and 2, while infants were listening to speech containing informative distributional or prosodic information respectively. Moreover, we performed a control experiment (Experiment 3), in which infants heard periods of structured speech, non structured speech and silence, and we looked for difference in the functional networks activation across conditions.

Chapter 4 studies memory constraints for the encoding of six-syllabic words in newborns. In Experiment 4 we tested if the first and last syllables were better encoded than middle syllables. In Experiment 5 we tested whether subtle pauses could enhance the encoding of intermediate syllables.

Chapter 5 explores how information is encoded when it is presented continuously. Specifically, we were interested in knowing if an abstract encoding of the segmented sequences is generated. We performed a series of behavioural experiments with adults and 5-month-old infants, using auditory and visual, and linguistic and non-linguistic stimuli, in order to explore for domain specificity.

Chapter 2

Speech Segmentation in neonates

2.1 The segmentation problem

In natural speech there are not pauses signalling the boundaries between words, and it is a still open question in language acquisition how infants do to identify the words of their language. We know adults use prosodic cues to extract many different properties of their own language. However, prosodic cues are usually assumed to be language specific, for example English words are usually stressed on the first syllable and French words in the last. If this is the case, how can infants learn these patterns before knowing the words and structure of their language?

A solution to this problem that has been proposed is the use of distributional cues. Infants could identify regularities based on the co-occurrence of elements in the speech without previous language knowledge, and use these regularities to extract segments as word candidates. Furthermore, the extracted information could be used to learn language specific cues like stress patterns.

The first evidence of the use of distributional cues in speech segmentation in infants came from the work of Saffran et al. (Saffran, Aslin, and Newport, 1996), in which they showed the 8-month-old were able to extract four three-syllabic words from continuous and flat synthesized speech after a brief familiarization. The regularities were based on the co-occurrence of syllables, and they formally described them in term of transitional probabilities (TPs), that is to say the frequency of a pair of syllables relative to the frequency of one of the syllables (the first in case of forward TPs, see equation 2.1; and the second in case of backward TPs, see equation 2.2). Other possible computations have been proposed —mutual information (Swingley, 2005), chunking (Perruchet and Vinter, 1998); but independently of what the brain computes, what it is crucial is that regularities are extracted only based on the distribution of syllables in the speech.

$$TP(x \rightarrow y) = \frac{freq(xy)}{freq(x)} \tag{2.1}$$

$$TP(y \rightarrow x) = \frac{freq(xy)}{freq(y)} \tag{2.2}$$

After Saffran and colleagues work extensive research has been done on the use of the distributional cues for segmentation. It has been extended to non-linguistic

stimuli like tones (Saffran et al., 1999), and to the visual domain (Kirkham et al., 2002; Bulf, Johnson, and Valenza, 2011). Moreover, non human primates (Hauser, Weiss, and Marcus, 2002) and even rats (Toro, Sinnett, and Soto-Faraco, 2005) have been shown to be sensitive to distributional cues over syllables. Statistical learning is a general learning mechanism, which works across domains and it is shared with many other species, which brings questions about its relevance for language acquisition: the fact that infants are sensitive to it does not necessarily imply that it is relevant during language acquisition. Moreover, it is possible to categorize, segment and predict upcoming events based on distributional properties, hence by referring to statistical learning in these terms we are not considering any specificity across domains or stimuli (Frost et al., 2015).

Different questionings have been exposed to the relevance of statistical learning for speech segmentation. Some authors have argued that when segmentation models based on distributional cues are applied to real infant direct speech, results are poor (Yang, 2004). In the same direction, another study shows that infants fail to extract the words when the complexity of the stimulus is increased by using words of different length (Johnson and Tyler, 2010). Other authors have questioned the relevance of TPs for speech segmentation by arguing that any sequence with consistent TPs is considered a word candidate even if it has never been heard (Endress and Mehler, 2009b). Another source of criticisms comes from studies combining distributional and prosodic cues. 8-month-old infants seem to rely more on stress than on distributional cues (Johnson and Jusczyk, 2001). Nevertheless, other study shows that 9-month-old do rely more on stress, whereas 7-month-old rely more on distributional cues (Thiessen and Saffran, 2003), favouring the hypothesis that distributional cues could be used to discover the stress pattern of the language. Finally, the extraction of words from continuous speech based on TPs seems to be constrained by prosodic contours in both adults (Shukla, Nespor, and Mehler, 2007) and 6-month-old (Shukla, White, and Aslin, 2011).

A second solution proposed to the problem of speech segmentation is the use of prosodic cues. The relevance of prosody for adult language processing is well documented. For example prosody has a hierarchical structure and adults can map this structure into syntactic components (Hayes, 1989; Nespor and Vogel, 2007; Langus et al., 2012). But also infants are very sensitive to prosody. Young infants, 6 to 9 month-old, react to disrupted and ill-formed prosodic units (Hirsh-Pasek et al., 1987; Nelson et al., 1989; Gerken, Jusczyk, and Mandel, 1994; Nazzi, Jusczyk, and Johnson, 2000; Soderstrom et al., 2003), and even neonates can recognize speech segments containing a word boundary from segments within words based on prosodic information (Christophe et al., 1994). These findings led to hypothesize that some aspects of prosody may be universal, thus infants could be sensitive to prosodic units and use them to segment speech in smaller chunks even without experience.

More evidence supporting the idea of universals in prosody, or in more general terms of perceptual biases that operate in speech perception, arrives from adult studies. Perceptual grouping biases have been described in adults. For example, adults prefer to group a sequence of tones with high initial pitch (Hayes, 1985; Hay and Diehl, 2007), a bias that seems to extend also to the visual domain (Peña, Bion, and Nespor, 2011). If this perceptual biases exist prosodic units, as intonational phrases, could be identified without previous experience and use for speech segmentation. In line with this hypothesis, Endress and Hauser, 2010 showed that adults can extract words from speech based on the prosody of unfamiliar languages.

From our perspective both types of cues are likely to be used during language acquisition, with potential switches in attention to one or the other during development. Statistical learning seems to be a very powerful mechanism that stands across domains and species. It seems not reasonable that such a strong learning mechanism will not operate for language acquisition. Nevertheless, the huge complexity in the input infants receive makes hard to believe that it is the only cue for speech segmentation in the first stages of language acquisition. Furthermore, the high sensitivity to prosody since birth, together with the existence of perceptual biases that can operate on speech, make prosodic cues a perfect candidate to restrict the space for statistical learning during speech segmentation.

In the current chapter we investigated neonates' capacities to use either distributional cues or prosodic boundaries, to segment and extract word candidates from continuous speech. Proving that infants are not only sensitive, but actually use these cues to extract information from speech since birth is crucial to evaluate the relevance of these two mechanism for speech segmentation. We registered cortical activity using fNIRS, which gives us also the possibility to understand the neural basis underling speech processing and statistical learning at birth.

2.1.1 Neuroimaging studies on statistical learning

Before going to the experiments I will summarize some results from previous neuroimaging studies on the use of distributional cues for segmentation.

The neural bases of statistical learning are still unknown, and indeed only a few studies have explored them. Moreover, the fact that it is a general mechanism opens the question of which computations occurs in domain specific neural circuits and which ones are shared across domains (see Frost et al., 2015 for a review).

In order to try to answer these questions some studies using fMRI have compared the activation towards structured versus random sequences of stimuli. Whereas activity in some areas seems to be modality specific, the activation of other regions does not seem to depend of the type of stimuli used. Visual stimuli elicit differential activation in higher order visual areas (Turk-Browne et al., 2009); and linguistic stimuli in higher auditory areas (McNealy, Mazziotta, and Dapretto, 2006;

Karuza et al., 2013; Cunillera et al., 2009), and pre-motor areas (Cunillera et al., 2009). Differently, the medial temporal lobe, including the hippocampus, activates independently of the modality (Turk-Browne et al., 2009; McNealy, Mazziotta, and Dapretto, 2006), denoting its role in generating a memory or representation. Further evidence of the engagement of the medial temporal lobe comes from the work of Shapiro et al. (Schapiro et al., 2014), who tested statistical learning across modalities in a patient with bilateral damage in this area and found no learning. These results denote that some areas are required independently of the modality —presumably areas involved in generating a representation— whereas other computations are performed by domain specific circuits. Interestingly, McNeally and colleagues found that participants ability to discriminate words from part-words correlated with the differential activation in the left superior temporal gyrus, evidencing the dominance of the left hemisphere for this task.

Other studies have used EEG to identify the neural correlates of statistical learning, and in particular its time course. After a couple o minutes of familiarization with a structured sequences of tones (Abla, Katahira, and Okanoya, 2008) or syllables (Cunillera et al., 2009) an ERP (N100, N400) appears locked to the first item (tone/ syllable) of the sequences constituting the stream. Interestingly, the ERP disappears with further exposition to the stream, pointing out that extracting the regularities is a very fast process. A bigger ERP for the first item of word-like sequences has been observed also in neonates using syllables (Teinonen et al., 2009) and tones (Kudo et al., 2011).

EEG has also been used to study how units during statistical learning are perceived by looking to differences in the power spectrum of the signal. When participants listen to a random sequence of syllables a peak in the power spectrum at the frequency corresponding to the syllables' duration appears. When participant listen to a structured language with long distance dependencies, a peak associated with the rule-words is observed, but only when brief pauses signal the edges between the words (Buiatti, Peña, and Dehaene-Lambertz, 2009). A similar result was also observed in 8-month-old pre-term and full-term infants (Kabdebon et al., 2015).

We studied speech segmentation in neonates using fNIRS. Notice that previous studies with newborns (Teinonen et al., 2009; Kudo et al., 2011) suggest that they are sensitive to distributional cues, but they do not show that neonates can actually segment and extract the words. In our study we investigate if after a familiarization with continuous speech, infants were able to recognize words from other sequences they had heard during the familiarization period.

2.2 Functional Near Infrared Spectroscopy (fNIRS)

Behavioural responses are hard to obtain with neonates , and appropriate imagining methods should be as less invasive as possible. In all the experiments with

newborns reported in this theses we used fNIRS. Here I summarize the most relevant aspects of the technique. For a general review see Ferrari and Quaresima, 2012; Scholkmann et al., 2014. For a review focus on infants studies see Minagawa-Kawai et al., 2008; Gervain et al., 2011; Lloyd-Fox, Blasi, and Elwell, 2010.

fNIRs uses near infra-red light to sense the blood saturation level in the cortex that is a measure of neural activity. Neural activity increases the oxygen consumption, and in response there is an increase in the local cerebral blood, the so called hemodynamic response. During a typical hemodynamic response for adults, an increase in the concentration of oxygenated haemoglobin (HbO_2) and a smaller decrease in de-oxygenated haemoglobin (Hb) is observed (see Figure 2.1). This is call the Blood-Oxygen Level Dependent signal (BOLD). The increase in HbO_2 is due to the fact that the oxygen transported to the area is higher to the oxygen used by the neurons, whereas the decrease in Hb is consequence of its flow to the veins. Human tissue is relatively transparent to near infra-red light, and its attenuation occurs mainly because of scattering and absorption by chromophores —primarily HbO_2 and Hb. Therefore, by emitting near infra-red light into the scalp at two different wavelength and by detecting the changes in intensity, HbO_2 and Hb concentration variation can be sensed (see Figure 2.1).

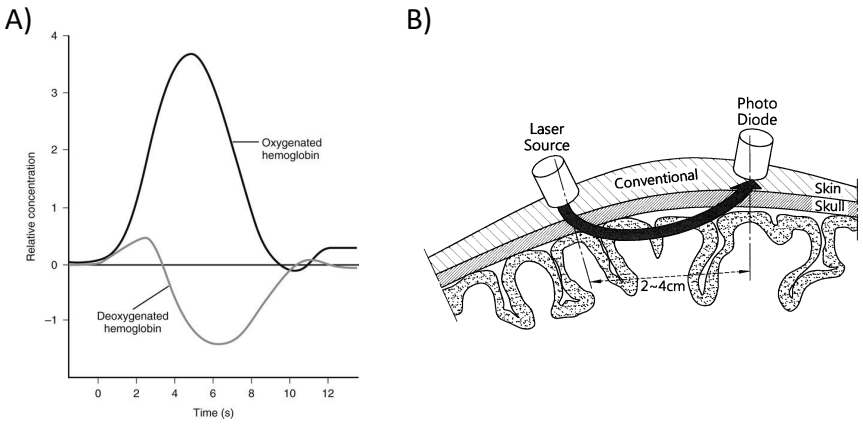


FIGURE 2.1: (A) BOLD signal. (B) Illustration of how fNIRS works.

fNIRS is a very suitable technique for infants studies for different reasons: 1. it is not invasive; 2. it is silent, which makes it particularly suitable for language experiments; 3. it is not so sensitive to movement as for example fMRI; 4. it potentially provides good localization, unlike EEG or MEG for which the sources are unknown 5. it has better temporal resolution than fMRI 6. it senses both HbO_2 and Hb. 7. it is relatively cheap and portable. However, it also present some disadvantages: 1. it can only measure cortical activity; 2. it does not provide anatomical information, which makes in practice harder to get good localization; 3. it is an indirect method and the HRF for infants is not well characterized, which can make

harder to interpret the results; 4. the temporal resolution is limited by the BOLD signal that is intrinsically slow, and it is much worse than in EEG and MEG. The fact that fNIRS is silent, non invasive and resistant to movement makes it a perfect method for language experiments with neonates. Moreover, infants skull is thinner and they have less hair than adults, providing a much better signal. Infants can be tested during sleep or quiet rest, and brain responses to auditory stimulation measured without requiring any active response.

Regarding infants' cognitive state during the experiment, a consideration has to be done. There is little research about how the state of sleep affects responses to auditory stimulation. Different studies have shown activity to language stimuli even in sleeping neonates (e.g.Peña et al., 2003; Gervain et al., 2008a; Benavides-Varela et al., 2011; Benavides-Varela et al., 2012; Mahmoudzadeh et al., 2013; Perani et al., 2011) and older infants (e.g. Homae et al., 2006; Homae et al., 2007; Nakano et al., 2009). Nevertheless, Dehaene-Lambertz, Dehaene, and Hertz-Pannier, 2002 found activation in right prefrontal regions in awake but not in sleeping 3-month-old infants —even if they did identify specific activation to speech in both, sleeping and awake infants. Although more specific research is needed to understand the effects of sleeping in speech processing during the first days of life; we can be confident, as demonstrated by numerous studies, that infants actively process speech even while sleeping. We tested neonates while sleeping or during resting state. We did not monitor the sleeping state, hence we cannot investigate for differences in this dimension.

In our experiments we used an habituation protocol. Habituation protocols are very common in behavioural experiments but they can also be implemented in neuroimaging studies. It is a well-documented phenomenon that repetition of the same stimulus leads to a decrease in neural activity and a subsequent change in the stimulus causes a recovery of neural activity, a pattern that is consistently found in neuroimaging studies across the lifespan from newborns and older infants (Dehaene-Lambertz et al., 2006a; Mahmoudzadeh et al., 2013) to adults (Dehaene-Lambertz et al., 2006b). In particular, the regions encoding the parameter that changed (Celsis et al., 1999), and frontal regions that are generally sensitive to novelty detection, show the greatest recovery of neural activity (Mahmoudzadeh et al., 2013; Nakano et al., 2009). Nakano et al., 2009 specifically tested novelty effects using fNIRS. Three-month old infants head a syllable (*pa*) repeated during a series of familiarization blocks, followed by a test block in which a new syllable (*ba*) was presented. They observed a broad recovery of the activation covering temporal and frontal regions (see Figure 2.2). Broader patterns of repetition suppression and recovery were also found in other fNIRS experiments (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012).

The experimental protocol we implemented takes advantage of these robust

habituation effects by examining differences in the recovery patterns between different types of test blocks. Stimuli presented during test blocks differ in a particular dimension from the familiarization stimulus. If the activity is recovered during a given test condition, this implies that infants learned something during the familiarization about that particular dimension. Instead if habituation effect are observed during the test period, it means that the changed parameter was not learned or encoded.

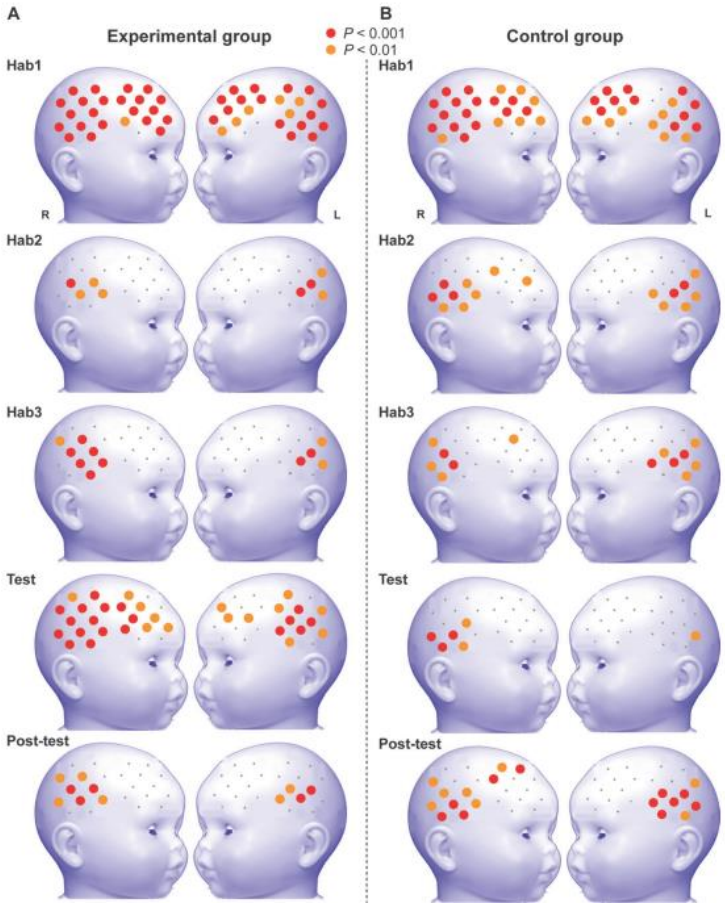


FIGURE 2.2: Habituation effect in fNIRS. The repetition of the same syllable (*pa*) produces a decrease in the HRF over frontal areas. When a new syllable is presented (*ba*) the signal is recovered. Adapted from Nakano et al., 2009.

2.3 Speech segmentation using distributional cues. Experiment 1

In a first experiment we asked if neonates are sensitive to distributional cues and can use this information to extract words from continuous speech. We familiarized infants with flat synthesized continuous speech with a structure analogue to

the stream used by Saffran, Newport, and Aslin, 1996, and afterwards we presented test blocks of either words or part-words, but now separated by silent periods. We used fNIRS to measure brain activation over fronto-temporal regions during the words and part-words test blocks. In order to distinguish words from part-words, newborns have not only to be sensitive to the co-occurrence of syllables during the familiarization phase, but also to form a memory or representation of the words. We expected that if infants segment and extract the words, we should observe an habituation effect for words and a novelty effect for part-words. Therefore a differential activation for the two types of sequences is a measures of recognition. In a typical HRF, HbO_2 increases and Hb decreases, thus if neonates segment end extract the words we expect HbO_2 to be higher for part-words than for words, whereas Hb should be higher for words.

2.3.1 Participants

All participants were healthy full-term neonates born to Italian-speaking mothers, with Apgar score ≥ 7 in the first minute and ≥ 8 in the fifth minute, diameter of head $\geq$ 33.0 cm, and no cefalhematomas. Experiment 1 included 40 participants (17 females; mean age 3.3 days; range 2-5 days except one 7 days-old infant; mean gestational age 39.1 weeks, rage 37-41 weeks; mean weight 3.306 Kg, SD 0.416 Kg) who provided data without motion artefacts from at least one of the test blocks per condition. Additional infants were tested but excluded from the final analyses because too many motion artefacts (n = 34), failure to complete the experiment due to fussiness (n = 10), a poor signal due to thick hair (n = 8), or due to technical problems (n = 2). This attrition rate is consistent with other studies using fNIRS with neonates (Gervain et al., 2011; Lloyd-Fox, Blasi, and Elwell, 2010). All newborns were recruited from the nursery at Hospital, Azienda Ospedaliera Santa Maria della Misericordia, in Udine, Italy. Parents provided informed consent. The Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

2.3.2 Stimuli

Stimuli were synthetized using the it4 Italian female voice of the MBROLA diphone database (Dutoit et al., 1996), with phoneme duration of 150 ms and a constant pitch of 200 Hz. Sequences were continuous with no pauses between syllables. The streams were built by concatenating in semi-random order, 4 three-syllables words (see Table 2.1). The syllables of the words had all a consonant-vowel structure, and were not repeated across words. Four Part-words were chosen to be tested in the experiment. Two of them were of the form $A_3B_1B_2$, and the other two of the form $A_2A_3B_1$, where A_i and B_i are the syllables of different words.

Stream	Words	Part-words
A	lamipe	mipedu
	duvoka	kanube
	nubefi	fitelu
	telugo	lugola
B	mipedu	peduvo
	vokanu	golami
	befite	nubefi
	lugola	telugo

TABLE 2.1: Stimuli for Experiment 1. Two familiarization streams and sets of Words and Part-words were used, and infants were randomly assigned to one or the other familiarization stream.

To synthesize the streams the only restrictions for the concatenation were that the same word could not appear twice in a row, and that two words could not alternate more than three times (the sequence $W_A - W_B - W_A - W_B$, where W_A and W_A are two words, was not allowed). As a result the streams had within words TPs of 1 and between words TPs of $\approx 1/3$. In particular for the long familiarization stream the average transitional probability between words was 0.3319 (SD = 0.0241, range [0.2833, 0.3667]). Each Word appeared 60 times and the Part-Words 17 to 21 times.

In order to avoid that results were driven by low level properties of the stimuli —that they were consequence of the specific set of words used— two different streams were created by using two sets of words (see Table 2.1). Noticed that half of the sequences that were Part-words for group A were Words for group B and vice-versa. Infants were randomly assigned to one or the other familiarization stream.

2.3.3 Procedure

The experiment started with a long familiarization and was followed by a set of four short familiarizations as reminders and test blocks (see Figure 2.3). Familiarization and test blocks were separated by periods of silence of random lengths between 25 and 30 s to allow the hemodynamic response to return to baseline. The duration for the long familiarization was 220 s, and for the short familiarizations 30 s. All familiarization blocks were ramped up and down during the first and last 6 s. The test blocks lasted 15 s and consisted of either the words or part-words repeated twice and separated by 0.5-1.5 s. The two kinds of test blocks were presented interleaved, and the order of presentation was randomized between participants —half of the infants heard first a Words test block and the other half heard first Part-words test block. The total duration of the experiment was 13 minutes.

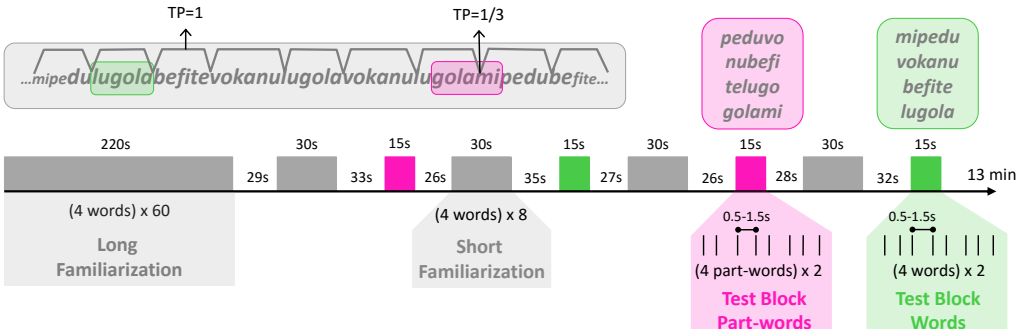


FIGURE 2.3: Schematic representation of the protocol for Experiment 1.The experiment started with a familiarization that lasted for 220 s, and was followed by a series of four short familiarizations and test blocks. During each test block 4 words or 4 part-words were presented separated by silences. Two of the test blocks contained words and the other two part-words and were presented interleaved. The order of presentation was counterbalanced across subjects.

2.3.4 Apparatus and data acquisition

Data were recorded using an ETG-4000 NIRS machine (Hitachi Medical Corporation, Tokyo, Japan) that uses two continuous light source wavelengths (695 and 830 nm). The separation between emitters and detectors was 3 cm, the sampling rate was 10 Hz, and total laser power output per fiber was 0.75 mW. Each probe consisted of nine fibers from which five were emitters and four were detectors (see Figure 2.4).

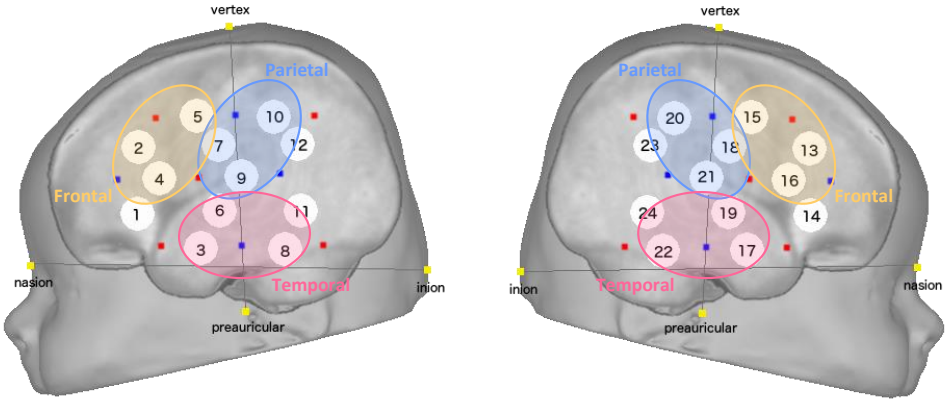


FIGURE 2.4: The recording was done using two silicon probes. Each probe contained 5 emitters (red squares) and 4 detectors (blue squares) with a 3 cm distance between them. In total each probe contained 12 recording points. The probes were positioned using skull landmarks, in particular the bottom detector was placed over the ear. The ROIs used for the mean activation analysis are shown.

The neonates were tested while lying in their cribs, asleep or in a state of quiet

rest, in a dimly lit sound-attenuated booth. Sound stimuli were presented at ≈ 60 dB via two loudspeakers placed on both sides at the feet of the infant's crib at a 30° angle. The speakers were connected to a Macintosh power PC G5 computer that simultaneously operated the NIRS machine and presented the auditory stimuli using PsyScope X software (Cohen et al., 1993). Both the NIRS machine and the computer were placed outside the experimental booth. Two silicon probes (each 7 cm x 9 cm), containing 12 recording points each, were used to keep the optical fibers in place. One probe was placed over the right side of the head (channels 1-12) and the other over the left (channels 13-24), using skull landmarks. Specifically, the bottom detector was placed above the ear and the probe was kept aligned along the anterior-posterior direction (see Figure 2.4 and 2.5). The positioning was chosen to maximize the recording from fronto-temporal regions. During the testing session an experimenter controlled the NIRS machine from outside the room, a second experimenter held the probes in place, a medical doctor blind to the experimental hypotheses assisted the neonate, and parents were free to remain in the booth or not. An infra-red video camera was used to monitor the infant's behaviour

FIGURE 2.5: Picture illustrating the testing procedure.

2.3.5 Data Analysis

The data pre-processing, calculation of the HRF, data rejection, and statistical analysis are described below.

Pre-processing. Data were pre-processed using custom functions and functions of the Homer2 NIRS package (Huppert et al., 2009), in MATLAB 2014b (MATLAB and Statistics Toolbox Release 2012b, n.d.), according to a general protocol used previously in other NIRS experiments (for a methodological review see Cooper et al., 2012; Gervain et al., 2011; for some applications see Ferry et al., 2016; Fekete et al., 2011; Peña et al., 2003). The pre-processing involved the following steps:

1. *Identification of periods of saturation or low signal to noise ratio.* We marked saturated samples as those with an intensity bigger than 5; and samples with low signal to noise ratio as those with a ratio between the power in the frequency bands 0.01-1Hz and 3-10 Hz smaller than 2.5 in a 30 s sliding time window. We pruned channels that showed saturation or low signal to noise ratio during more than 50 % of the time at any of the two wavelengths, and in a later stage of the pre-processing we spatially interpolated the signal of these channels (see step 5).

2. *Conversion of the intensity to optical density, and linear detrending.* We converted the intensity to optical density using the *hmrIntensity2OD* function of the Homer2 NIRS package. Afterwards we linear detrended the entire time series.
$$OD = -log\left(\frac{I}{mean(I)}\right) \tag{2.3}$$

3. *Detection of motion artefacts by fast changes in the signal.* We used a modified version of the *hmrMotionArtfactsByChannel* function of the Homer2 NIRS package. The function identifies samples in a moving time window of length tMotion $\pm$ tMask as motion artefacts if there is a change greater than a threshold, amp_thresh, or if any value in the time window is bigger than std_thresh standard deviations. To take into account that the signal for different subjects and channels can differ in its variability, we modified the function to set the threshold independently for each channel and subject. The distribution of the maximum absolute change for all the time windows was calculated and those windows containing changes that deviated from the first or third quartiles in more than amp_thresh times the interquartile range were considered as having motion artefacts. In particular, if Δ_i is the maximum absolute change for the time window i, and q_1 and q_3, are the first and third quartiles, then samples in a time window are marked as having motion artefacts if $\Delta_i < q_1 - amp\_thresh \cdot (q_3 - q_1)$ or $\Delta_i > q_3 + amp\_thresh \cdot (q_3 - q_1)$. The parameters we used were: tMotion = 1s, tMask = $\pm$ 0.5s, amp_thresh = 1.75 and std_thresh = 20. Because motion artefacts should affect all channels, we considered there was a motion artefact if fast changes were detected in at least 6 out of the 24 channels at any of the two wavelengths.

4. *Correction of motion artefacts by target PCA.* We used the *hmrMotionCorrectPCA* of the Homer2 NIRS package, and we removed 0.97 of the variation of the data (Cooper et al., 2012; Yucel et al., 2014).

5. *Spatial interpolation of pruned channels and periods with fast changes in less than 6 channels.* To do so we used modified Shepard's spatial interpolation, with a maximum radius, R, of 4.5 cm. This implies that only neighbour channels

closer than 4.5 cm contributed to the interpolation. If x_j is the signal of channel i that has to be interpolated then:

$$x_j(t) = \frac{\sum_{i=1}^{N} w_i \cdot x_i(t)}{\sum_{i=1}^{N} w_i} \tag{2.4}$$

with,

$$w_i = \left(\frac{max\Big(0, R - d(x_i, x_j)\Big)}{R \cdot d(x_i, x_j)} \right)^2 \tag{2.5}$$

6. *Calculation of the relative changes in HbO_2 and Hb.* To do so we used the *hmrOD2Conc* function of the Homer2 NIRS package, that uses the modified Beer-Lambert law. In this modification of the classical Beer-Lambert law a single factor is added in order to correct for two phenomenons: the longer effective distance travel by light due to the scattering through the tissue, and for the assumption that only a fraction of the volume is responsible of light absorption (Huppert et al., 2009; Strangman, Franceschini, and Boas, 2003; Cope and Delpy, 1988).

$$\Delta OD = L_{ij}^{\lambda} \cdot DPE^{\lambda} \cdot (\varepsilon_{Hb}^{\lambda} \cdot [\Delta Hb] + \varepsilon_{HbO_2}^{\lambda} \cdot [\Delta HbO_2]\,) \tag{2.6}$$

L_{ij}^{λ} is the distance between emitter and detector, DPE^{λ} is the differential path length factor, and $\varepsilon_{Hb}^{\lambda}$ and $\varepsilon_{HbO_2}^{\lambda}$ are the attenuation coefficients. Because two wavelength are used the changes in HbO_2 and Hb can be estimated.

7. *Re-detection of motion artefacts.* This step is necessary because some motion artefacts —usually too long or too strong— are not properly corrected. We used the same parameters than in step 3, and again we considered that a motion artifact was present when fast changes were detected in at least 6 out of 24 channels.

HRFs calculation. First, we band-pass filter the pre-processed data between 0.01 Hz and 0.50 Hz, in order to reduce slow systemic physiological hemodynamic fluctuations, mainly respiratory signals (0.2-0.4 Hz) and blood pressure changes (0.08-0.12 Hz); fast cardiac oscillations (≈1 Hz), and high-frequency instrument noise. Second, we extracted the HRFs from the HbO_2 and Hb time series for each Test Block. To account for the delay of the BOLD response we cut from -5s to +35s from the onset of the stimuli. We used the mean value in the period [-5s-0s] and [30s-35s] to calculate a linear baseline trend that was removed from the signal. Third, after data rejection, we calculated an average HRF per infant per channel per condition, that we used for statistical analysis.

Data rejection. HRFs for blocks in which motion artefacts were present and not properly corrected were rejected. We visually inspected the data that appeared as outliers but that had not been identified as motion artifact by the algorithm, and we compared to annotations made during the experiment and the video recording. If motion artefacts were identified, data was manually rejected. Infants were included in the analysis only if they contributed at least with one good test blocks per condition.

Statistical Analysis. In traditional statistical analysis either the peak or the mean value for HbO_2 and Hb during the HRF is calculated and used for the analysis. This approach, even if still nowadays extensively used, presents some disadvantages. First, the temporal information of the HRF is lost, which is a potential issue because the time course of the activity is not necessarily the same over different cortical areas. Second, the mean or peak value of the HRF has to be inferred. If the mean value is used, a time window where the average is calculated has to be arbitrarily chosen; whereas when the peak value is used the peak has to be estimated, which is not a trivial process that can be strongly affected by artefacts. Third, because the majority of the NIRS systems have multiple recording points, the problem of multiple comparisons should be considered, issue that has not always been appropriately addressed (e.g. Gervain et al., 2008a). One traditional approach is to run independent t-test per channel and correct the p-values per multiple comparisons. Conventional corrections as Bonferroni work fine with a small number of channels, but become over conservative with more recording points, specially since the activity is not independent across channels. More appropriate methods control for false discovery rate and take into account the dependency of the data points (for a review see Singh and Dan, 2006; for some implementations see e.g. Benavides-Varela et al., 2011; Nakano et al., 2009). Another traditional statistical analysis approach consist in defining ROIs and performing an ANOVA with ROIs and experimental conditions as factors (e.g. Yang et al., 2016; Sato et al., 2013; Peña et al., 2003). This method overcomes the problem of multiple comparisons by a priori defining the ROIs, with the consequent lost of spacial information.

In summary all traditional approaches imply a reduction of the data set according to a priori decisions from the experimenter. Recently implemented cluster based non-parametric analysis, instead, allows researches to control for multiple comparisons without requiring a reduction of the data set and a priori decisions, thus retaining power for more detailed analysis.

We decided to perform two different statistical analysis: a cluster based permutation analysis and traditional mean activation analysis over different ROIs.

Cluster Based Permutation Analysis. This analysis was initially developed for EEG and implemented in the Fieldtrip Matlab toolbox (Oostenveld et al., 2011)

and has previously been applied to NIRS data (e.g. Ferry et al., 2016; Edwards et al., 2015; Mahmoudzadeh et al., 2013). When the time courses of a signal under different condition have to be compared without reducing the dimensionality of the data set, two main problems are faced. First, the number of comparisons is huge —one for every time bin and every registration point. Second, data points are not independent. The cluster based permutation analysis is a non-parametric test that identifies significant differences between conditions in spatio-temporal clusters, retaining sufficient statistical power while adequately controlling for the problem of multiple comparisons. The method is designed to account for the fact that similar responses are expected between samples that are close in time or space, thus non independent.

Briefly, the steps of the analysis are the following. First, t-tests are conducted between each pair of sets of data points, $x_A(channel, time)$ and $x_B(channle, time)$), where x_A and x_B contain the data for all subjects in the two conditions. Second, cluster candidates are identified by grouping all temporally and spatially adjacent pairs with a p-value smaller than a chosen threshold (standard value of .05). It has to be noticed that this threshold does not affect the false discovery rate, its only purpose is to identify cluster candidates. Third, cluster-level statistics are calculated for each cluster candidate by summing the t-values of the t-test for every comparison included in a cluster. The cluster-level statistics are larger for larger clusters (more t-values are added) and for clusters with larger differences between conditions (larger individual t-values are added). Finally, a permutation analysis is used to evaluate whether the cluster-level statistic is significantly different from chance. A null distribution is obtained by randomizing the conditions and the proportion of random partitions that produce a cluster-level statistic greater than the observed statistic is the Monte Carlo p-value for the cluster.

We used the Cluster Based Permutation Analysis to compare the HRFs for test blocks of Words and Part-words in the interval [0s, +30s] from the onset. Before running the analysis we smooth the data by down sampling to 1 Hz, which does not imply a loss of temporal resolution because the BOLD signal is intrinsically slow (≈10 s) (Mahmoudzadeh et al., 2013). As a result we obtained 744 pairs of data points to compare (24 channels x 31 samples). We performed two tails one sample t-tests and we used 0.05 as threshold p-value to select the pairs of samples to build the clusters. We considered two pairs of samples temporally adjacent if they were consecutive (time difference of 1s), and spatially adjacent if they were at a distance < 3 cm (See Table 2.2). We ran 1000 randomizations to obtain the Monte Carlo p-value.

Mean Activation Analysis. We calculated the mean activation in the period [+10s, +30s] from the onset of the stimuli in four ROIs per hemisphere. The regions for the left hemisphere were: frontal (channels 2, 4 and 5), temporal (channels 3, 6 and 8), and parietal (channels 7, 9 and 10); and analogue ones for the right hemisphere

Left Probe		Right Probe	
Channel	Neighbours	Channel	Neighbours
1	2 3 4	14	13 17 16
2	1 4 5	13	14 16 15
3	1 6	17	14 19
4	1 2 6 7	16	14 13 19
5	2 7	15	13 18
6	3 4 7 8 9	19	17 16 18 22 21
7	4 5 6 9 10	18	16 15 19 21 20
8	6 11	22	19 24
9	6 7 11 12	21	19 18 24 23
10	7 12	20	18 23
11	8 9 12	24	22 21 23
12	9 10 11	23	21 22 24

TABLE 2.2: Neighbour channels used for the cluster based permutation analysis.

(see Figure 2.4). We compared the activity for Words and Part-words across the different regions using a 3-ways ANOVA with within subjects factors condition (Words/ Part-Words), region (frontal/ temporal/ parietal) and hemisphere (left/ right).

2.3.6 Results

Results for both methods of analysis are presented below. Figures and Tables for results based on Hb are presented in the Appendix A.

Cluster Based Permutation Analysis. We found a broad increase in the hemodinamic response for Part-words and a decrease for Words. HbO_2 concentration was higher for Part-words than Words in two clusters. One comprising all channels in the right hemisphere within the time window [9s - 26s] ($P_{HbO_2}cluster_1 < 0.01$), and one in the left hemisphere including all channels except for channel 10 again within the time window [9s - 26s] ($P_{HbO_2}cluster_2 < 0.05$) (see Figure 2.6.a). For Hb we found the opposite patter of results, the activity was higher for Words than Part-words. The analysis revealed one cluster over the right hemisphere, including all channels except for channel 14 and within the time window [11s - 31s] ($P_{Hb}cluster_1 < 0.01$); and one cluster including all channels of the left hemisphere, within the time window [17s - 29s] ($P_{Hb}cluster_2 < 0.05$) (see Figure A.1.b).

Mean Activation Analysis. The results were consistent with the results obtained by the cluster based permutation analysis. For HbO_2 the 3-ways ANOVA revealed a main effect of condition (Words / Part-words) ($F(1,78) = 10.337$, $P < 0.005$), but not effect of hemisphere or region ($P > 0.05$), and not significant interactions (P

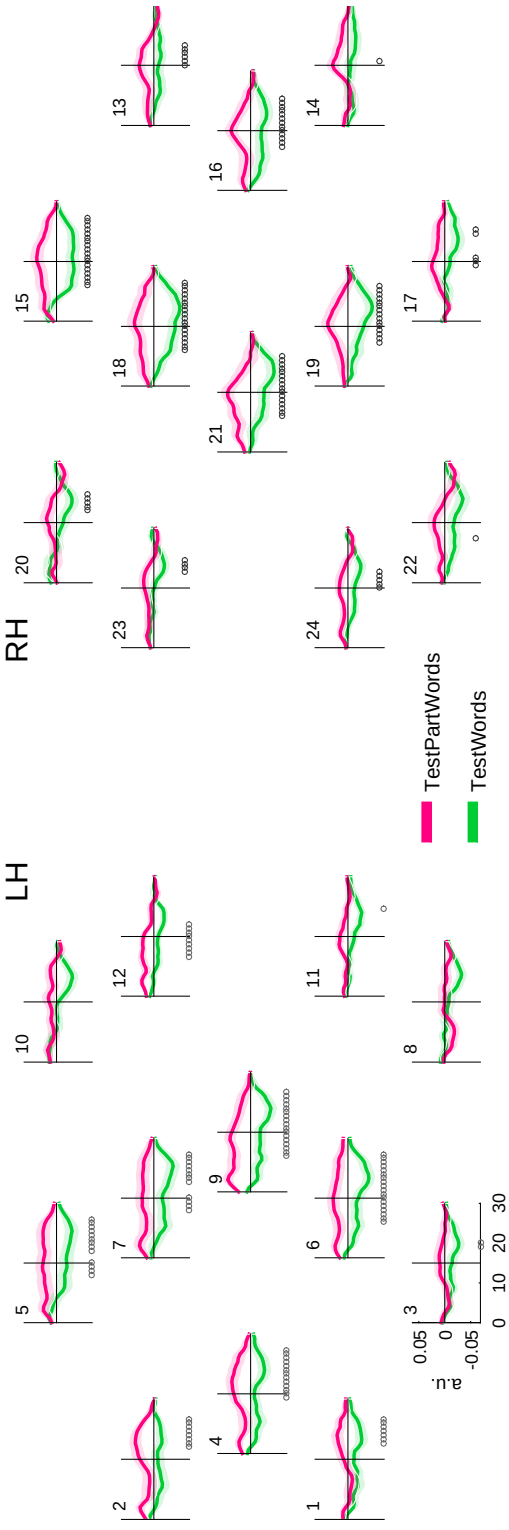


FIGURE 2.6: Cluster based permutation analysis for Experiment 1 using HbO_2. HRFs for Words (green) and Part-words (pink) during test blocks. Vertical lines indicate the onset and offset of the stimulus. Marks below the individual channels show the time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

> 0.05) (see Figure 2.7 and Table 2.3). For Hb we found a main effect of condition (Words / Part-words) (F(1,78) = 6.6219, P < 0.05), and again not effect of hemisphere or region (P > 0.05), and not significant interactions (P > 0.05) (see Figure A.2, and Table A.1).

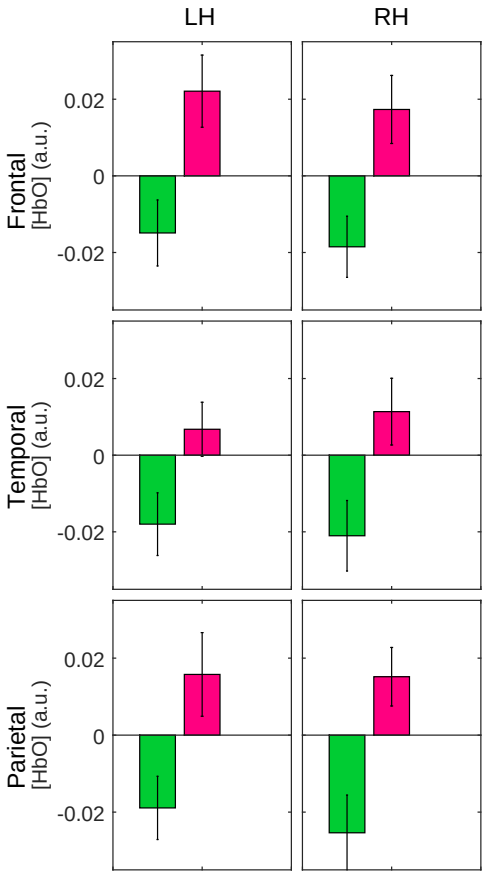


FIGURE 2.7: Mean activation analysis for Experiment 1 using HbO_2. The mean activity for Words (green) and Part-words (pink) over each of the three regions of interest in the left and right hemisphere is shown. Error bars represent standard errors

2.3.7 Discussions

In this experiment we found that after less than 4 minutes of listening to an artificial language neonates showed a higher response for part-words and words over fronto-temporal regions, providing the first evidence to our knowledge of word extraction from continuous speech by neonates.

Previous studies with neonates had revealed sensitivity to distributional cues over syllables, tones and visual shapes, but did not provide direct evidence of parsing and extraction of the sequences. Two EEG experiments, one using syllables (Teinonen et al., 2009) and one using tones (Kudo et al., 2011), show that after some minutes of familiarization with a continuous stream the ERPs towards the first, second and third items of the three-items sequences conforming the

HbO_2	SumSq	DF	MeanSq	F	P
Hemisphere	0.51269	39	0.01315		
Error (Hemisphere)	0.00064	1	0.00064	0.62566	0.43370
Region	0.03999	39	0.00103		
Error (Region)	0.00385	2	0.00193	1.56699	0.21520
Condition	0.09586	78	0.00123		
Error	0.14037	1	0.14037	10.33745	0.00260
Hemisphere:Region	0.52958	39	0.01358		
Error (Hemisphere: Region)	0.00058	2	0.00029	0.43091	0.65150
Hemisphere: Condition	0.05291	78	0.00068		
Error (Hemisphere: Condition)	0.00051	1	0.00051	0.42443	0.51860
Region: Condition	0.04653	39	0.00119		
Error (Region: Condition)	0.00193	2	0.00096	0.71688	0.49150
Hemisphere: Region: Condition	0.10481	78	0.00134		

TABLE 2.3: Statistical analysis for Experiment 1. 3-ways ANOVA using HbO_2 as dependent measure. Hemisphere (left/ right), region (frontal/ temporal/ parietal) and condition (Edge/ Internal) are within subject factors.

stream differ. Further evidence of neonates sensitivity to distributional information comes from a behavioural experiment by Bulf et al. (Bulf, Johnson, and Valenza, 2011), in which they report longer looking time to a random sequence of shapes than towards a previously learned sequence containing structure. But results from our experiment go beyond. After a period of familiarization, brain activation toward words and part-words in isolation was different, meaning that neonates segmented and extracted the items based in certain regularities. Infants could have been sensitive to the co-occurrence of syllables, or they could have perceived the stream conformed by distinct chunks; but independently of the underlying computation, they extracted and retained the regularity even when the familiarization was over, and they applied the extracted pattern to words in isolation.

What neonates are actually computing in our study deserves some discussion. Both, words and part-words, were conformed by syllables present in the familiarization, and more over both types of sequences appeared in the stream, but words were ≈ 3 times more frequent. Consequently, our result can have different interpretations. One possibility is that neonates are sensitive to the co-occurrence of pairs of syllables, TPs. A second option is that they are tracking the frequency of the different three-syllabic sequences. Based on our results we cannot discriminate between these two options, but research with older infants indicates that segmentation is done based on TPs (Aslin, Saffran, and Newport, 1998). We consider unprovable a change in such a basic mechanism during development and we hypothesize neonates are tracking the same distributional property that older infants. The specific computational mechanism underlying statistical learning and segmentation remains unclear, and it is source of intense debate for both infants and adults (e.g. Perruchet and Pacton, 2006; Conway and Christiansen, 2001; Frost et al., 2015; Dehaene et al., 2015), but this discussion is beyond the aims of our experiment.

Statistical learning is a general mechanisms, hence we can wonder if neonates learning capacities based on distributional information differ across domains. Older infants are able to extract word like sequences from a continuous streams of tones (Saffran et al., 1999) and visual shapes (Fiser and Aslin, 2002; Kirkham et al., 2002), and as we have already mentioned neonates seem to be sensitive to distributional regularities also in other modalities (Kudo et al., 2011; Bulf, Johnson, and Valenza, 2011). Nevertheless, we cannot infer from newborns sensitivity to distributional information that they are able to extract units in other modalities. Extracting the sequences involves high memory demands, thus even if newborns are able to perform the computations differences in learning performance across domains may arise from differences in neonates' capacities to encode different types of information. It remains as an interesting question if newborns' ability to extract sequences from speech overcomes their capacity in non linguistic domains.

What can we say about the observed patter of activation towards words and part-words? We tested the recognition of the words taking advantage of the strong habituation effects observed in fNIRS (Nakano et al., 2009; Benavides-Varela et al., 2011). The origin of the habituation effect is not clear, therefore the straightforward interpretation of areas showing activation or deactivation being involved in processing specific aspects of the stimuli may be a bit adventurous. In their work with 3-month-old Nakano et al., 2009 hypothesized that a distributed network involving sensory, associative and prefrontal areas may be responsible. They found prefrontal regions responding to novelty, whereas temporal areas showed a maintained response that was however modulated in amplitude by the number of repetitions of the stimulus. In our experiment we found a broad activation for the novel stimulus (part-words) and deactivation towards the familiar one (words), and we did not identify any significant effect of localization in the mean activation analysis using ROIs. But this should not be interpreted as a global activation over all cortical areas. Our recordings were mainly limited to frontal and temporal regions, because we intentionally located the probes where we expected bigger differential responses. Despite this limitation in the recorded areas, the hemodynamic response was stronger in frontal than in more temporal and posterior channels. In fact, the cluster based permutation analysis showed differences over posterior and temporal channels that were more restricted in time than over frontal channels, evidencing a stronger effect in frontal regions.

A further limitation for identifying spacial effects in our results is the poor spacial localization of the channels. The probes were positioned based on skull landmarks, but the size and shape of neonates' heads are highly variable. In addition testing infants usually requires that probes positioning is a fast and smooth process, impeding high accuracy, which adds more variability to the data. Together, this high between subjects variability in the localization of the channels, plus an effect that is intrinsically broad, entail a limitation on the identification

of the exact localization and extension of the effect. In order to have a comprehensive picture of the habituation effect, cortical activity should be recoded from more areas and better methods for identifying the exact localization of the channels in the cortex should be developed.

Despite the limitation just discussed, habituation effects in fNIRS seem to be a suitable measure to test recognition in neonates. For example, word segmentation from continuous speech has been extensively studied in numerous studies and under different conditions in older infants using behavioural methods (e.g. Saffran, Newport, and Aslin, 1996; Thiessen and Saffran, 2003; Johnson and Tyler, 2010). However, behavioural responses are hard to get in neonates. Our results demonstrate that fNIRS provides a recognition measure with at least the value of a behavioural measure, that can also provide some insights of the underlying cortical processes.

To summarize, this experiment is the first prove that neonates are not only sensitive to distributional cues in the speech, but that they are actually able to extract words out of it, evidencing amazing learning and memory capacities since very early in life.

2.4 Speech segmentation using prosodic cues. Experiment 2

In the previous experiment we showed that even at birth infants are able to extract words from continuous speech when only distributional cues are present. However, the stimuli used in Experiment 1 were flat and natural language is not. Phonemes change in pitch, intensity and duration, conforming a very rich prosody, and numerous studies have shown infants sensitivity to it (Hirsh-Pasek et al., 1987; Nelson et al., 1989; Gerken, Jusczyk, and Mandel, 1994; Nazzi, Jusczyk, and Johnson, 2000; Soderstrom et al., 2003; Christophe et al., 1994).

In the current experiment we asked if Italian neonates could extract words from continuous speech based on prosodic cues alone. To test this we used the exact same procedure than in Experiment 1. The only difference between experiments was during the familiarization phases: during the familiarization of the current experiment no informative distributional cues were present, and the only cue for segmentation were prosodic contour of four Italian phrase. Crucially, words and part-words during test blocks were presented flat as in Experiment 1. This implies that a differential responses towards words and part-words cannot be attribute to the detection of well or ill-formed prosodic units. Getting a novelty effect towards part-words and, specially, an habituation effect toward words, indicates that neonates are able to extract the words neglecting irrelevant aspects as supra-segmental information (changes in pitch and duration of the phonemes).

Furthermore, we will explore the presence of hemisphere differences in the activation during the recognition of the words. A superiority of the right hemisphere for the processing of prosody was observed using fNIRS in 4-yeas-old kids (Wartenburger et al., 2007) and even 3-month-old infants (Homae et al., 2006). We wonder if this effect will appear during the presentation of the words and part-words even if they are presented flat.

2.4.1 Participants

All participants were healthy fullterm neonates born to Italian-speaking mothers, with Apgar score ≥ 7 in the first minute and ≥ 8 in the fifth minute, diameter of head ≥ 33.0 cm, and no cefalhematomas. The experiment included 40 participants (19 males; mean age 3.2 days; range 1-5 days; mean gestational age 39.2 weeks, rage 37-41 weeks; mean weight 3.400 Kg, SD 0.459 Kg) who provided data without motion artefacts from at least one of the test blocks per condition. Additional infants were tested but excluded from the final analyses because too many motion artefacts (n = 28), failure to complete the experiment due to fussiness (n = 10), a poor signal due to thick hair (n = 14), or due to technical problems (n = 1). All newborns were recruited from the nursery at Hospital, Azienda Ospedaliera Santa Maria della Misericordia, in Udine, Italy. Parents provided informed consent. The Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

2.4.2 Stimuli

We used two familiarization stream and sets of Words and Part-words that were the same than in Experiment 1 (see Tables 2.1). Infants were randomly assigned to one or the other group. Stimuli were synthesized using the it4 Italian female voice of the MBROLA diphone database (Dutoit et al., 1996) and sequences were continuous with no pauses between syllables. Unlike Experiment 1 phonemes had variable duration and pitch, and the distribution of syllables in the familiarization stream was different. We built the streams by concatenating the words in a fix order, resulting in uniform TPs of 1 and an equal frequency of Words and Part-words, meaning that the distribution of syllables was uninformative for the segmentation task. Instead prosodic information was added to the streams, in particular variations in pitch an duration. To do so we recorded four Italian sentences of the form CVCVCV (see Table 2.4) from a female native Italian speaker, we extracted the pitch and duration of each phoneme, and we imposed them to the words in the stream (see Figure 2.8). To over-impose the pitch and duration, we first fitted the pitch using the *SmoothingSpline* function of MATLAB, and calculated the pitch in 12 equidistant points per phoneme. Afterwards, we normalize the pitch of each phrase to a mean pitch of 200Hz, and the total duration

to 900 ms conserving the relative duration of each phoneme. Finally, we concatenated the contours in a fix order and we used them to synthesize the streams with MBROLA. Note that because both, the order of words and the prosodic contours are fixed, the pitch and duration for each phoneme are conserved along the entire familiarization. In other words, each word was marked by a specific prosodic contour.

Italian Phrases
ti pare?
si vede!
lo faro!
come va?

TABLE 2.4: Italian phrases recorded to extract the prosodic contours that were added to the familiarization stream of Experiment 2.

The Words and Part-words used for the test blocks were flat, with a constant pitch of 200 Hz and phoneme duration of 150 ms, exactly as in Experiment 1.

2.4.3 Procedure

The presentation of the stimuli, duration of familiarizations, test blocks, and silence periods, was the same than for Experiment 1 (see 2.3.3). All familiarization blocks were ramped up and down during the first and last 6 s. The two kinds of test blocks were presented interleaved, and the order of presentation was randomized between participants. In brief, the only difference between Experiment 1 and 2 are the familiarization blocks (see Figure 2.9).

2.4.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

2.4.5 Data Analysis

Idem than for Experiment 1. See 2.3.5.

2.4.6 Results

Results for both methods of analysis are presented below. Figures and Tables for results based on Hb are presented in the Appendix A.

Cluster Based Permutation Analysis. For HbO_2 we found a higher activation for Part-words than Words in one cluster in the right hemisphere including all channels besides channel 13, within the time window [11s - 28s] ($P_{HbO_2} cluster_1 < 0.01$). We did not find any significant cluster in the left hemisphere (see Figure 2.10). For

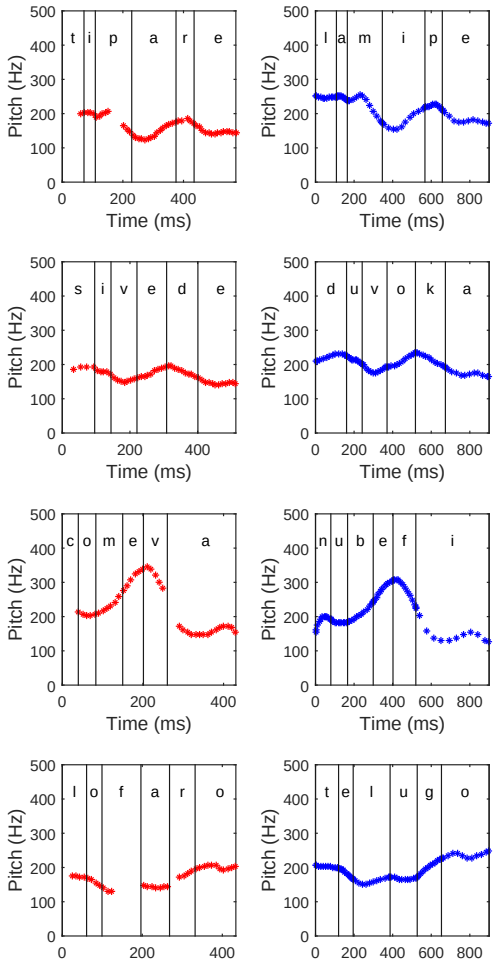


FIGURE 2.8: Prosodic contours for Experiment 2. On the left the real prosodic contours. On the rights the prosodic contours extracted and imposed to the words of the stream.

Hb we found a significant cluster in the right hemisphere including all channels besides channel 13, within the time window [1s - 8s]. Activity was higher for part-words then for words ($P_{Hb}cluster_1 < 0.05$) (see Figure A.3.b).

Mean Activation Analysis. The results confirmed what we obtained by the cluster based permutation analysis. For HbO_2 the 3-ways ANOVA revealed a main effect of condition (Words / Part-words) ($F(1,78) = 8.9235$, $P < 0.005$), but not effect of hemisphere or region ($P > 0.05$), and a marginally significant hemisphere x condition interactions ($F(1,39) = 4.1108$, $P = 0.0495$) (see Figure 2.11). Post-hoc multiple comparison analysis using Turkey-Kramer correction showed that the interaction was due to a differential activity in the right and left hemisphere only significant for part-words ($P < 0.05$) (see Table 2.6). For Hb we found no main effects, and non interactions ($P > 0.05$) (see Figure A.4).

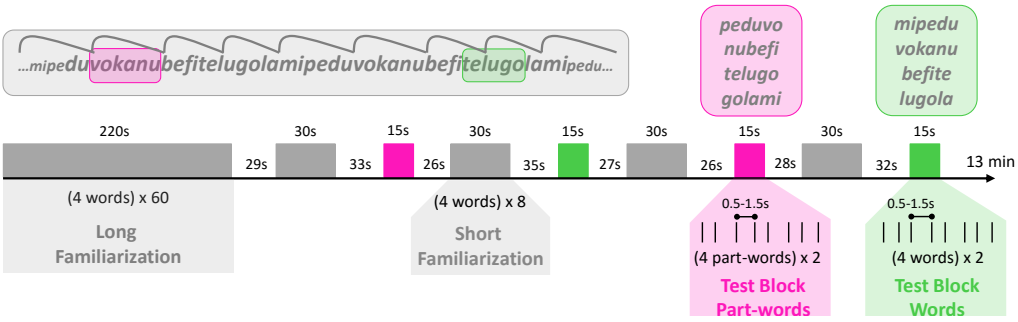


FIGURE 2.9: Schematic representation of the protocol for Experiment 2. The experiment started with a familiarization that lasted for 220 s, and was followed by a series of four short familiarizations and test blocks. During each test block 4 words or 4 part-words were presented flat and separated by silences. Two of the test blocks contained words and the other two part-words and were presented interleaved. The order of presentation was counterbalanced across subjects.

HbO_2	SumSq	DF	MeanSq	F	P
Hemisphere	0.48109	39	0.01234		
Error (Hemisphere)	0.00193	1	0.00193	2.04675	0.16050
Region	0.03684	39	0.00094		
Error (Region)	0.00348	2	0.00174	2.07758	0.13210
Condition	0.06526	78	0.00084		
Error (Condition)	0.09746	1	0.09746	8.92347	0.00480
Hemisphere: Region	0.42596	39	0.01092		
Error (Hemisphere: Region)	0.00076	2	0.00038	0.72497	0.48760
Hemisphere: Condition	0.04092	78	0.00052		
Error (Hemisphere: Condition)	0.00509	1	0.00509	4.11084	0.04950
Region: Condition	0.04833	39	0.00124		
Error (Region: Condition)	0.00035	2	0.00017	0.17766	0.83760
Hemisphere: Region: Condition	0.07655	78	0.00098		

TABLE 2.5: Statistical analysis for Experiment 2. 3-ways ANOVA using HbO_2 as dependent measure. Hemisphere (left/ right), region (frontal/ temporal/ parietal) and condition (Edge/ Internal) are within subject factors.

	Difference	StdErr	P	CI lower	CI upper
LH: Part-words - Words	0.02198	0.00947	0.02552	0.00284	0.04110
RH: Part-words - Words	0.03501	0.01063	0.00212	0.01350	0.05650
Part-words: LH - RH	-0.01053	0.00419	0.01617	-0.01900	-0.00210
Words: LH - RH	0.00250	0.00434	0.56785	-0.00628	0.01130

TABLE 2.6: Statistical analysis for Experiment 2 using HbO_2. Post-hoc analysis using Tukey-Kramer correction.

2.4.7 Discussions

In the current experiment we found that after being familiarized with continuous synthesized speech with Italian intonational phrases over-imposed, neonates

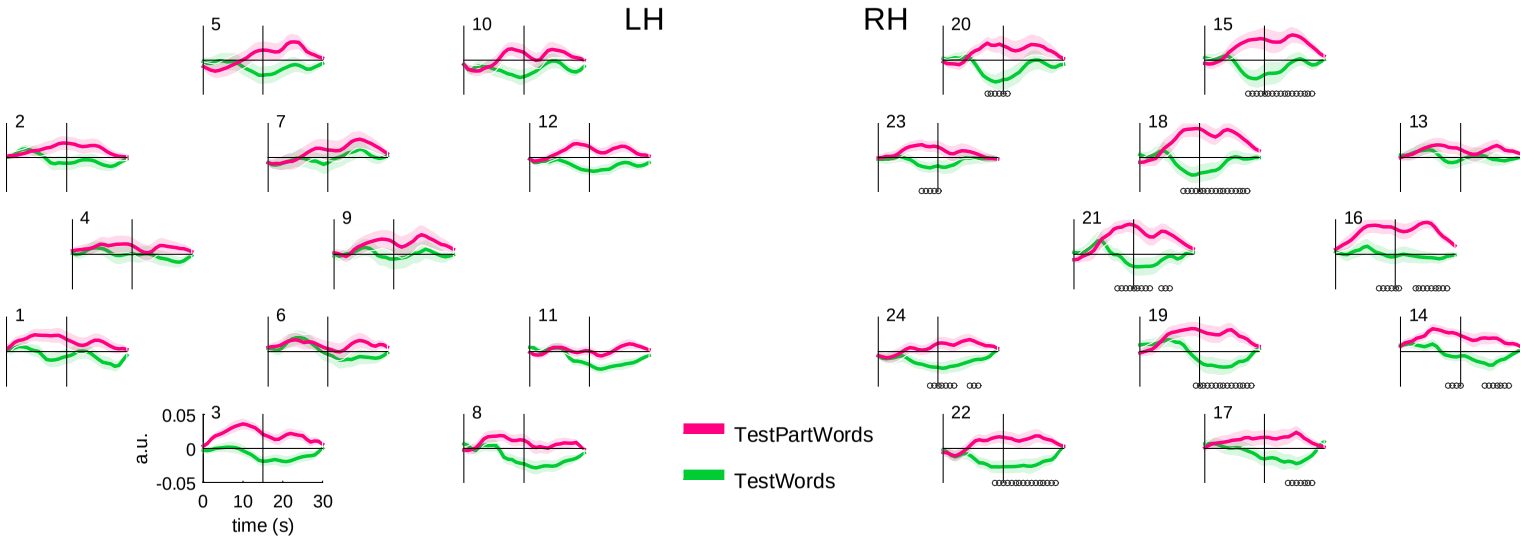


FIGURE 2.10: Cluster based permutation analysis for Experiment 2 using HbO_2. HRFs for Words (green) and Part-words (pink) during test blocks. Vertical lines indicate the onset and offset of the stimulus. Marks below the individual channels show the time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

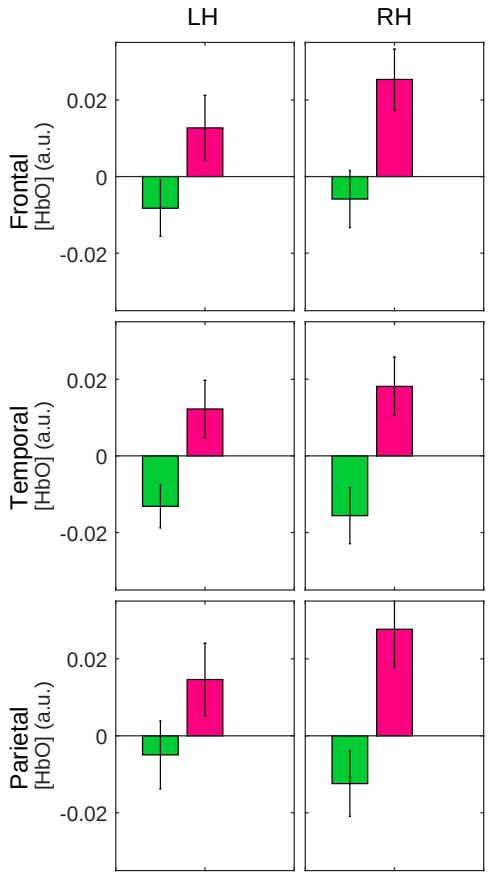


FIGURE 2.11: Mean activation analysis for Experiment 2 using HbO_2. The mean activity for Words (green) and Part-words (pink) over each of the three regions of interest in the left and right hemisphere is shown. Error bars represent standard errors

show higher hemodynamic responses for sequence of syllables that were straddling the intonational phrase (Part-words) than for those inside the prosodic units (Words).

Infants sensitivity to prosody has been showed in numerous studies, but our results show more than that. We provided evidence that neonates can use the prosodic contours of their mother language to segment and extract words from continuous speech, even if it is the only available source of information. To our knowledge this is the first evidence of this capacity in infants. Previous studies were either segmentation studies with older infants, in which both distributional and prosodic cues were present (Johnson and Jusczyk, 2001; Thiessen and Saffran, 2003; Shukla, White, and Aslin, 2011); or were focus on the distinction of well and ill-formed prosodic units (e.g. Hirsh-Pasek et al., 1987; Christophe et al., 1994). We used a stream built by concatenating the syllables in a fixed order, thus transitional probabilities were uniform and uninformative, and the frequency of words and part-words was exactly the same. Consequently, the only possible explanation of our results is that neonates parse the speech based on the prosodic

contours.

A remarkable aspect of our findings is that neonates did no limit to recognize a well form prosodic unit, but they actually extracted the words and recognize them from part-words even if presented flat. Sequences during test blocks did not share the prosodic information with the words in the familiarization stream —the duration and pitch of the phonemes was different. Although a word presented flat is acoustically pretty different than the same word with its intonational phrase, we did observe an habituation effect. This implies, not only that infants parsed the speech, but also that they formed a representation of the words that goes beyond superficial acoustical aspects.

Neonates' capacity to use prosodic contours to segments speech, implies that the prosodic contours are perceptual units for them, which leads to an important question: is it because of innate biases or is it a consequence of pre-birth experience with Italian prosody? Numerous studies demonstrate that even before birth infants actively process auditory stimuli. Fetuses are able to remember linguistic stimuli (DeCasper and Spence, 1986) and music (Kisilevsky et al., 2004) with which they were presented during the last trimester of gestation. Newborns prefer their mother's voice over other voices (DeCasper and Fifer, 1980; Querleu et al., 1984), and they are able to recognize their mother tongue language from languages belonging to different rhythmic classes (Mehler et al., 1988; Nazzi, Bertoncini, and Mehler, 1998). On the other hand there is evidence of the existence of perceptual biases operating in speech processing of phonemes (Eimas et al., 1971; Streeter, 1976), and of grouping biases in adults (Hayes, 1985; Hay and Diehl, 2007) and infants (Yoshida et al., 2010), hence it is also a possibility that this ability is independent of previous exposure. Unfortunately we cannot answer this question based on our results. In order to do so the same experiment should be tested using intonational phrase from a language belonging to a very different rhythmic classes.

There are two main considerations I would like to make about the pattern of cortical activation we observed. First, the cluster based permutation analysis returned significant results only in the right hemisphere. Consistently with this, the ANOVA on the mean activation showed a marginally significant interaction condition x hemisphere, driven by a higher activation for part-words in the right than in the left hemisphere. This hemispheric difference could be attributed to a right hemispheric dominance for the perception of prosody, which has been observed in different studies in adults (Friederici and Alter, 2004; Zatorre and Belin, 2001; Meyer et al., 2002; Gandour et al., 2004), 4-years-old (Wartenburger et al., 2007) and infants (Homae et al., 2006; Homae et al., 2007; Telkemeyer et al., 2009). Extracting the words in the current task requires integrating auditory information over longer periods of time than Experiment 1, for which the success on the task relies much more on the integration of segmental information. A right hemisphere dominance for prosody processing is a plausible explanation

for the stronger differential response for part-words and words over right than left frontal regions. Nevertheless, we have to be prudent with the conclusions. Habituation and novelty effects are slightly stronger on the right hemisphere (Nakano et al., 2009), therefore it could be also the case that the effect reveals weak in the left hemisphere as a consequence of this.

A second consideration regards the robust effect observed using HbO_2, in contrast with the less consistent effect based on Hb. The cluster based permutation analysis revealed that the Hb concentration change was higher for part-words than words right after the onset of the test block, which goes in the opposite direction of what expected —we expected a decrease in Hb for the novel stimulus and an increase for the familiar one. However, the cluster appears very close to the onset of the stimulus. In a typical BOLD signal Hb increases at the very beginning, hence it may be that what we are observing is this early effect. Moreover, this could make harder to obtain significant results later in time, because responses from different subjects could have different latencies. Notwithstanding it remains as an open question why we observed this early effect, our general conclusions are not affected, we know that HbO_2 is in general more reliable and results based on it were clear.

In sum, we can say that Italian neonates are able to use Italian intonational phrases to extract words from continuous speech even when distributional cues are not informative. Moreover, they seem to create a representation of the words that neglects some acoustical aspects of the signal. This result sheds light on the role of prosody in early language acquisition. Infant-directed speech has been shown to exaggerate prosodic markings and contain longer pauses (Fernald, 1989; Fernald, 1992). These cues have been shown to help older infants attend to and segment speech (Johnson and Jusczyk, 2001; Thiessen, Hill, and Saffran, 2005) but our results suggest that infant-directed speech may facilitate language acquisition even during the first days of life. The arguably universal use of infant-directed speech (Bryant and Barrett, 2007; Fernald, 1992) may result from these very early uses of prosodic boundaries to begin to segment and encode speech.

2.5 Chapter Discussions

Our two experiment show impressive learning and memory capacities of the neonatal brain. Experiment 1 revealed that neonates are sensitive to distributional regularities in the speech and that they rely on them to segment and extract words from continuous stream. Experiment 2 showed that infants perceive prosodic contours as units and that they are able to extract words contained in them. This findings have important implications on the problem of language acquisition. Since birth infants seem to be ready to start extracting statistical regularities from language and to use prosodic information to parse speech. Furthermore, our results

suggest that neonates form short term representation of words that are independent of low level acoustic properties.

While the main goal of the experiments was to test neonates segmentation abilities, infants have to form some representation of the words in order to recognize them. Thus the experiments also shed light on how infants encode word forms. The recognition of words and part-words in Experiment 1 could be done independently of which aspects of the speech are encoded. However, for Experiment 2 the recognition task becomes harder. Infants need to favour the encoding of the segmental information that characterizes each phoneme over supra-segmental information as pitch and duration. This affirmation may seem to contradict previous studies showing that infants during the first month of life focus more on supra-segmental information, but this is not necessary the case. Even if young infants seem to be very sensitive to prosody —it provides not only linguistic information, but it is also crucial for social interaction— supra-segmental information may be discarded during the encoding of segmental information. The idea is consistent with experiments reporting that young infants (Jusczyl, Pisoni, and Mullennix, 1992; Kuhl, 1983) and even pre-terms (Mahmoudzadeh et al., 2013) can disregard voice differences in phonemes recognition. The results are also compatible with the most widely accepted theory of phonemes perceptions. Even if infants at birth are be able to discriminate all possible phonemes contrasts and they became tune to the phonemes of their own language during the first years of life (e.g. Polka and Werker, 1994; Werker and Tees, 2002; Werker et al., 1981; Maye, Werker, and Gerken, 2002), infants have a categorical perception of phonemes since birth (Eimas et al., 1971; Eimas and Miller, 1980; Dehaene-Lambertz and Pena, 2001). In sum, our results have indirect implication on word representation and phoneme perception by showing that neonates can create quiet sophisticate short term representations of words.

Notwithstanding the result suggest amazing capacities in neonates to start extracting regularities from speech, the extrapolation to the problem of language acquisition has to be done carefully. The stimuli we used were far from being ecological: we used synthesize speech with very simple distributional and prosodic properties. In natural language both, phonemes distribution and supra-segmental properties are much more complex, and furthermore both are present together. But the fact that they appear together does not necessarily make the task harder. As I discussed in the introduction, we can distinguish two main issues in the segmentation problem. First, the real relevance of pure distributional cues in natural language; and second, if universal in prosody exist or if they are entirely language specific. We cannot answer any of these questions based on our results, but I will argue that to shed some light to the bootstrapping problem we should not focus on the two problems in isolation.

Regarding the first problem —the relevance of distributional cues— natural speech is much more complex in its distributional properties than the artificial

languages used in experiments, but at the same time it is not flat. One possibility is that statistical learning does not work by integrating distributional information over long periods of time, but that it is a more local process. In fact, for example, words straddling prosodic contours are not recognized by adults (Shukla, Nespor, and Mehler, 2007). Perceptual units determined by prosodic components could restrict the computations related with the occurrence of syllables.

But what about the origin of these perceptual units: Is prosody specific for each language? If so, how can it be used for segmentation? Our second experiment demonstrated that independently if because of universal perceptual biases, or as consequence of the extraction of acoustical regularities from pre-birth experience, infants are able to create perceptual units before having enough experience with language at the segmental level. How the computation of statistical regularities across phonemes interacts with supra-segmental prosody, and how this interaction may change along development appears as a rich field of research.

To finalized, our results also contribute in the understanding of functional brain development. In both experiments we found a strong activity in frontal areas for novel linguistic stimuli. This contradicts the classical view of frontal areas being too immature and not connected with the language network at birth, and adds to recent finding showing activity in this region in neonates (Benavides-Varela et al., 2011) and even in pre-terms (Doria et al., 2010; Mahmoudzadeh et al., 2013). In addition, in Experiment 2, we observed stronger habituation/novelty effects in the right hemisphere, potentially as consequence of the superiority of the right hemisphere in processing supra-segmental information. If that is the case, it means that this advantage of the right hemisphere translates into a better recognition of the words even when the supra-segmental information is not there any more and it has to be discarded.

Chapter 3

Functional connectivity in neonates

In the present chapter I will present results from a new functional connectivity analysis we performed on fNIRS recordings from neonates. In order to contextualized our research I will first make a brief introduction to brain connectivity and in particular to functional connectivity.

3.1 Brain connectivity

The goal of brain connectivity is to understand the interactions between different structures and the physiological process that give place to function. In other words, to uncover the structural and functional organization underling the huge number of functions and cognitive process that the brain is able to perform (for general reviews see Petersen and Sporns, 2015; Park and Friston, 2013). When we talk about connectivity the focus is on the links between the different elements of the system. The elements defining our system will depend of the level at which we are looking to the brain, from a molecular level, to a neural level, arriving to brain region. Independently of the level at which we are working, we can refer to these elements and their connections as a network.

The development in the last years of neuroimaging techniques that enable to register brain activity over different areas —fundamentally fMRI— leaded to the emergence of the field of functional brain connectivity. In functional connectivity the links between different areas are not based on their structural connections, but on their coordinate activity.

3.1.1 A formal description of networks: Graph theory

Graph theory is the mathematical study of networks, where a network is a set of nodes (vertices) and links between the nodes (edges). One important benefit of graph theory is that it enables to quantitatively characterize the topological properties of a network using a small number of representative measures.

The brain is a very complex network in which some distinctive features can be identified at different levels. For example the brain presents functional specialization, cortical hierarchy and zones of convergence. These properties arise from some key organizational principles that are shared with other complex systems, and result into networks that are neither random nor order, that tend to have highly connected hubs, and to be organize in modules; properties that are easily characterized by a relatively small number of graph measures. This is why in the last years graph theory started to be extensible applied to the study of anatomical and functional brain connectivity (see Bullmore, Sporns, and Solla, 2009 for a general review, Rubinov and Sporns, 2010 for a review on graph measures, and Bassett and Gazzaniga, 2011 for a theoretical discussion).

In anatomical networks a connections is for example a white matter track, whereas in functional networks is a temporal association between the activity in two areas. The first step for building a networks is to determine its nodes or vertices, and afterwards to estimate the associations between all of them, thus to obtain the association matrix that links all the vertices. Graphs can be binary, meaning that vertices are or are not connected; or weighted, when the value of the edge denotes the strength of the association. Moreover links can be undirected, when there is not distinction between the two associated nodes; or directed, when there is directionality from one node to the other. In order to build a functional network usually the association matrix is thresholded, and only the stronger connections are kept. In the case of unweighted networks the matrix is also binarized (see Figure 3.1). The most used measures of association for undirected functional networks are coherence, correlations or mutual information, and for directed networks Granger causality can be used.

Many complex systems including different biological networks present similar properties. It is commonly assumed that these properties arise from shared building principles mainly involving a minimization of the wiring cost, while maximizing the capacity of integrating and segregating information. In a network the wiring cost can be measured by the *density* of the network, meaning the number of edges respect to the total number of possible links. *Segregation*, in the context of network science, refers to an organization in densely interconnected groups of nodes; whereas *integration* is the facility to merge information from different regions. A network highly segregated without loss in integration present what is call *small-world* properties, that from theoretical consideration it can be considered as a balance between efficiency and wiring cost (Bullmore and Sporns, 2012). From a functional perspective small-worldness implies the presence of segregated neural processes, with a the same time a good capacity to share information between them.

Numerous graph measures are used to described brain networks, predominantly describing of the above mentioned properties. I will now focus on some basic measures that we will use in our analysis.

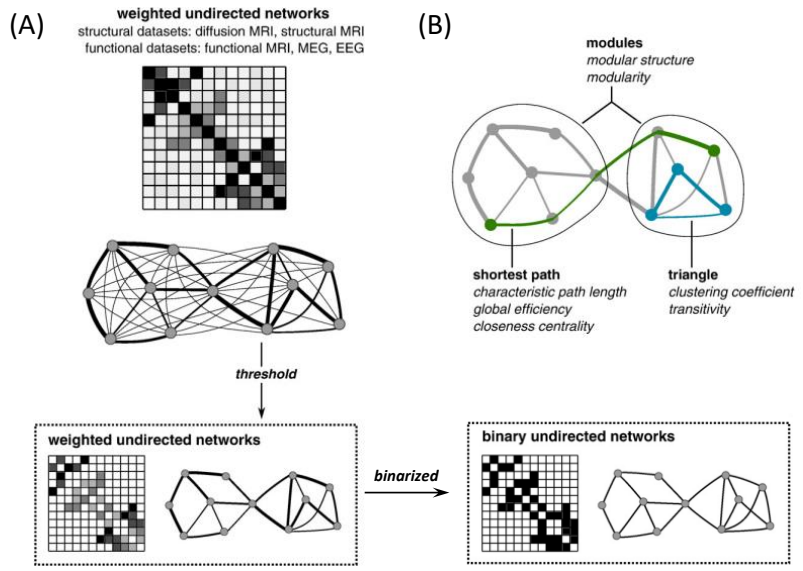


FIGURE 3.1: (A) Illustration of the steps for the construction of functional or structural networks. (B) Illustration of some network measures. Adapted from Rubinov and Sporns, 2010

.

A first fundamental graph measure is the *degree* of a node. The degree of a node is the number of links that connect it with other nodes in the net, and the distributions of all degrees in the network is a distinctive property of it. In particularly complex networks have distributions with long tales, because some nodes have more connections.

Segregation is measured in terms of clustering or groups of interconnected nodes. The *cluster coefficient* of a node is the fraction of its neighbours that are also connected to each other, and ranges from 0 to 1. The mean cluster coefficient of the network is the mean across all the vertices of the network.

Measures of integration are based on the paths, meaning the links necessary to arrive from one vertices to another. The *characteristic path length* of a network is the average of the shortest path length between all pairs of nodes. If a node is disconnected the path length for nodes involving it are defined as infinite, therefore the characteristic path length is not well defined for disconnected networks. Instead the global efficiency can be used. *Efficiency* is defined as the inverse of the shortest path length, thus it is zero for disconnected nodes.

Small-world is the balance between integration and segregation, and is defined as the ration between the mean cluster coefficient and the characteristic path length. A network with small-word properties presents different hubs interconnected with each others, which implies the existence of central nodes interconnecting the hubs. A measure of centrality is *betweenness centrality*, defined as the proportion of all the shortest path of the network that pass though a node.

The application of graph theory to neural networks in the last years has enabled to show that the brain presents a modular and hierarchical organization, in both its structure and functional architecture, which makes possible a good integration of information in and between local and segregated specialized circuits (He and Evans, 2010).

3.1.2 Structural and functional networks and how anatomy constrains function

The interpretation of results relative to structural networks is much more straightforward than for functional networks. In structural networks edges are real anatomical links between brain areas, whereas in functional networks edges are a measure of the synchronization of the activity. One general approach to the study of functional connectivity is the recording of spontaneous activity during rest (see Power, Schlaggar, and Petersen, 2014 for a review). The origin of this activity is unknown, but resting state functional connectivity has shown across different studies very reproducible networks (e.g.Smith et al., 2009). Furthermore, these networks seem to be persistent across variations in the resting state —like close and open eyes (McAvoy et al., 2012), or light anaesthesia and sleep (Larson-Prior et al., 2011)— and even across tasks (Cole et al., 2014). Some examples of well described functional networks are the default mode network, a fronto-parietal network, a dorsal and a ventral attention system, a sensorimotor visual and auditory systems (Power et al., 2011).

A big question in neuroscience is the link between structure and function (Lichtman and Denk, 2011) and the relation between structural and anatomical networks is an instance of this inquiry. Structural networks will necessary constrained the space of functional networks, which is evidenced by the similarities that are observed between functional and anatomical networks; nevertheless, it will not fully determine them. On one hand two regions that are anatomical connected do not necessary have to show a synchronized activity; and on the other hand areas that do not present direct anatomical links are many time functionally connected (Honey et al., 2009; Hermundstad et al., 2013). Functional connectivity does not represent direct anatomical connections, but it is, the result of all the potential existing connections (Adachi et al., 2012), and reflects a functional relation.

Structural and functional connectivity differ in their nature. While structural networks are relatively fixed (at least within certain time and space scales), functional networks are dynamic, which settles the bases for cognition (Park and Friston, 2013). If the functional networks we build really represent brain areas that are functionally associated, we expect differences in functional connectivity to emerge under different mind states or tasks. Even if the main aspects of functional networks seem to reflect anatomical connections, and even more, structure can be infer from functional networks (Hermundstad et al., 2013); evidence of the dynamic nature of functional connectivity has been found. Modifications in

functional architecture associated with different brain states have been observed during different tasks (Cole et al., 2014; Davison et al., 2015; Hermundstad et al., 2013), learning (Bassett et al., 2011; Bassett et al., 2015), deep sleep (Horovitz et al., 2009) and deep anaesthesia (Heine et al., 2012; Boveroux, Vanhaudenhuyse, and Phillips, 2010; Barttfeld et al., 2014). Furthermore, different psychiatric disorders have been associated with differences in functional connectivity (e.g. Leonardi et al., 2013; Assaf et al., 2010; Lynall et al., 2010; Jones et al., 2012).

3.1.3 Dynamic functional connectivity

Till recently functional connectivity had been studied by measuring the synchronized activity of brain areas during long periods of time. This approach assumes stationarity, or in other words a constant interdependence of the signals along the recording period. This assumption is not strictly correct. Functional connectivity is a dynamic process, thus studies assuming stationarity make a description that is a time average, which may explain why functional connectivity resembles so much structural connectivity (Deco, Jirsa, and McIntosh, 2011).

A complete characterization of the functional architecture of the brain means describing its complexity in both the spatial and temporal dimensions. An effort to achieve a temporal description of the electrical activity (for example in EEG, MEG studies) exists from long time, but only recently an interest for describing the dynamic of the functional connectivity in terms of BOLD signal arose (see Hutchison et al., 2013a), which constitutes a growing field of research. The general approach to study functional connectivity dynamics is based on estimating connectivity in sliding time windows, follow by divers subsequent steps. For example brain states have been identified by clustering, either based directly on the correlation matrices (Allen et al., 2014) or on topological features of the networks (Bassett and Gazzaniga, 2011). Meanwhile other works focus on the identification of common activation patters (Majeed et al., 2011).

The research in the field is quite new and functional connectivity dynamics is far from been understood, but some main observations have been done in the last years. First, functional connectivity shows fluctuations in magnitude denoting that brain passes from more to less efficient connectivity states (Hutchison et al., 2013b; Zalesky et al., 2014). Second, the brain seems to move between a set of discrete network states (Allen et al., 2014). Finally, some connections seem to be more variable than others: frontal regions —and more generally areas linking different modules— show higher flexibility (Cole et al., 2013), and their variability has been associated with learning and task performance (Braun et al., 2015; Bassett et al., 2011).

Despite dynamic functional connectivity appears as a promising field, it also presents some methodological issues: 1. Dynamic functional connectivity estimation is highly sensitive to physiological noise (Chang et al., 2013), and head motion (Power et al., 2012; Yan et al., 2013). 2. If correlations are used, the level

of correlation due to stochastic noise should be estimated (Handwerker et al., 2012). 3. The levels of no neural noise can change along the experiment, modifying the functional connectivity. 4. The same region can be involved in different networks —networks overlap— hence methods allowing to identify and separate these networks would be ideal. 5. Clustering techniques commonly used to identify networks states are not so robust. 6. The best size of the time window for sliding time window techniques is also an open question. The window has to be short enough to capture transient changes, but it has to be long enough to capture the slow BOLD fluctuations and to have a sufficient signal to noise ratio. Usually time windows of 30-60 seconds have been used for fMRI data (Hutchison et al., 2013a).

3.1.4 Brain connectivity in infants

Studying the functional organization of the brain during development gives us the possibility of understanding which aspects of the functional architecture in adults are present from birth and which ones change during development, either by the interaction with the environment, or due to the developmental time course.

The literature about functional connectivity in infants is much reduced than for adults, but some studies suggest that functional networks similar to adult networks are present from birth (Fransson et al., 2007; Doria et al., 2010). Doria and colleagues showed that the development of the functional networks observed in adults occurs in the third trimester of gestation. They measured BOLD signal at rest for infants between 29 and 43 weeks of gestational age and identified visual, auditory, somato-sensory, motor, default mode, fronto-parietal and executive control networks. The visual and auditory networks were stable even in pre-term infants, whereas the somato-sensory, motor, default mode, and executive control networks increase in connectivity and stabilized during the last trimester. This study suggests that networks at birth are not limited to local areas but include also associative areas. Despite this, there is also evidence that cortical network are more restricted to primary sensory and motor areas than in the adult brain (Fransson et al., 2011).

Some other studies have investigated resting state functional connectivity using fNIRS. Homae et al. (Homae et al., 2010; Homae, 2014) measured functional connectivity in neonates, 3 and 6 month-old infants, and found that: 1. Correlations increase during development: 2. Homologous connections were in general stronger than other connections. 3. For neonates, between the homologous connections, frontal connections were stronger than parietal, temporal and occipital connections. 4. At 3-month-old homologous temporal connections were still weaker than the rest, but no differences were found at 6-month-old. The authors hypothesize that this pattern during the development of homologous connections may be a consequence of the development of the corpus callous from front to back.

3.2 Functional connectivity while listening a structured sequence of syllables. Experiment 1

In Experiment 1 we proved that neonates are able to segment continuous speech based on distributional cues and to recognize the words from the part-words after familiarization. Evidence of this comes from the habituation effect for words and novel effect for part-words we observed during test trials. These effects extend broadly over temporal and frontal areas, and are consequence of the repetition of a stimuli and its recognition, but do not reflect the learning process. In order to explore which neural process are behind the statistical learning observed in Experiment 1, we performed a dynamic functional connectivity analysis during the familiarization phase.

It has to be noted that to our knowledge, no previous studies have explored dynamic functional connectivity neither in infants, nor using fNIRS in general. Moreover, little work has been done in general on functional connectivity using fNIRS with infants (Homae et al., 2010; Homae, 2014; White et al., 2012) and adults (Lu et al., 2010; Zhang et al., 2010; Zhang et al., 2011; Zhang et al., 2014; Sasai et al., 2011; Niu et al., 2012; Medvedev, 2014; Molavi et al., 2013). Hence, our results will add to the little literature that there is in the field.

The main aims of the analysis are: 1. to search for differences in strength of the connections between different areas or types of connections, meaning to describe the connectivity; 2. to investigate for functional connectivity changes along the task; 3. to identify functional networks active during the task based on connectivity changes; 4. to see if either some aspects, or changes along the familiarization, of the functional connectivity are correlated with infants recognition of the words and part-words (learning). A description of the functional architecture during the familiarization will reflect both, aspects that are proper of the task, and properties that would appear also during resting state and may be consequence of the anatomical connections. In this respect, it is interesting to try to identify which aspects are intrinsically associated with the task by correlating them with task performance. Furthermore, because a learning process is going on during the familiarization, it is compelling to try to describe the dynamic of the connectivity. Learning is not only associated with stationary properties of the networks, but also depends of its dynamic (Bassett et al., 2011; Bassett et al., 2015).

With these considerations in mind we decided to performed a dynamic functional connectivity analysis by estimating the functional connectivity in sliding time windows. To described the connectivity we took two different approaches.

In a first approach we simply described the strength and time variability of connections between specific regions (e.g. left and right intra-hemispheric connections, homologous and non-homologous inter-hemispheric connections).

In a second approach, we built functional networks and described them using graph measures. We were particularly interested on the dynamic, hence we wanted a method that identifies networks not only based on which areas have more synchronize activity, but that also reflects which areas change their connectivity together. In order to identify the more representative modes of variation of our data set we decided to apply principal component analysis (PCA) to the fluctuations in connectivity. Specifically, we implemented a method adapted from Leonardi and colleagues work (Leonardi et al., 2013). In their study they applied PCA to a connectivity dynamic obtained from resting state fMRI data of normal subjects and multiple sclerosis patient. In our case we applied PCA to the fNIRS data of newborns during the segmentation task, and based on it we built functional networks. Afterwards, we first described the topological properties of the networks using graph measures; and second, we looked for correlations between the activity of these networks and task performance. A detail description of the analysis is presented below.

3.2.1 Participants

Idem than for Experiment 1. See 2.3.1.

3.2.2 Stimuli

Idem than for Experiment 1. See 2.3.2.

3.2.3 Procedure

Idem than for Experiment 1. See 2.3.3.

3.2.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

3.2.5 Data Analysis

A complete description of the analysis is presented bellowed.

Pre-procesing. We performed the same steps that were described for Experiment 1 (see 2.3.5). Because connectivity analysis are highly sensitive to motion artifacts —strong signal correlation— they should be corrected or removed from the time series. In the pre-processing described for Experiment 1, we corrected motions artifacts (step 4); nevertheless, the data was not always well reconstructed. In order to improve the quality of the data and eliminate fast changes, we applied target PCA a second time to data segments in which fast changes were re-detected

(step 7 of the pre-processing, see 2.3.5). We finally band-pass filtered the data between 0.01 Hz and 0.10 Hz, as it has been done in previous fNIRS connectivity analysis (Niu and He, 2014).

Data rejection. We excluded subjects that presented motion artifacts during more than a 30% percent of the duration of the familiarization, or with more than 6 rejected channels. Out of the 40 subjects of Experiment 1 10 were rejected because of too many motion artifacts, the other 30 were included in the analysis.

Estimation of the functional connectivity. To estimate the connectivity between pairs of channels, we used Pearson correlation coefficients. We computed the correlation coefficients in a sliding time windows of length 30s (300 samples) and step 2 s, obtaining a sequence of 96, 24×24 correlation matrices per subject (see Figure 3.2). In order to have normally distributed correlation coefficients, we Fisher transformed the correlation matrices.

Strength and variability of the connections. We first vectorized the Fisher transformed correlation matrices by taking the upper part of each matrix, and we obtained a vector of length 276, $276 = (24 \times 24 - 24)/2$, per time window per subject, which represents the time course of the strength of each connection. Second, we calculated the mean and the standard deviation of each correlation coefficient along the time dimension. Afterwards, with the purpose of investigating topological differences in the strength and variability of the connections across participants, we defined four asymmetry indexes. To calculate the indexes we first identified different types of connections (see Figure 3.2), and afterwards we computed the average strength and variability across them.

We defined the indexes based on previous findings describing: 1. anatomical asymmetries during development between hemispheres (Glasel et al., 2011; Leroy et al., 2011); 2. strong functional connections for homologous regions (Homae et al., 2010; Homae, 2014; Perani et al., 2011); 3. the presence of local functional networks (Doria et al., 2010; Fransson et al., 2011); 4. and intra-hemispheric fibre tracks that are not yet developed in the neonatal brain, as the dorsal pathway that connects the temporal cortex with the inferior frontal gyrus (Perani et al., 2011).

The indexes are:

$I_{Intra-Inter}$ Based on intra and inter-hemispheric connection pairs. Intra hemispheric connections were connection within either the left or the right hemisphere; whereas inter-hemispheric connections were the connection between the left and right hemisphere. This index allows us to check global differences between the within and between hemispheres connectivity.

$I_{Hom-nonHom}$ Based on homologous and non-homologous inter-hemispheric connection pairs. A connection was identified as homologous if the connected points were in opposite hemispheres and the distance from the symmetric

point to the connected channel was smaller or equal than 3 cm. With this index we can explore if, as in previous findings, homologous regions are functionality more related than non homologous.

$I_{Right-Left}$ Based on left and right intra-hemispheric connections pairs. Left connection were all the connections within the left hemisphere; whereas right connections were all the connections within the right hemisphere. This index compares the functional connectivity within the right and left hemisphere. Because anatomical asymmetries have been observed they may reflect in functional differences.

$I_{Short-Long}$ Based on short and long intra-hemispheric connections pairs. A short connection was a connection between two channels closer than 3 cm. The functional organization of infant's brain is more local compared with the adults brain, probably because some anatomical connections between distant regions are not developed yet. With this index we aim to check for this.

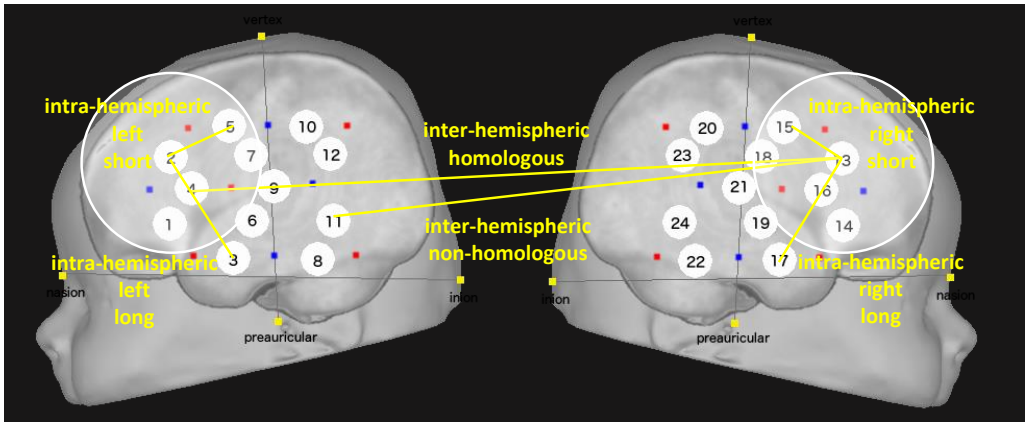


FIGURE 3.2: Illustration of the different types of connections according to the regions involved. The circle around channels 3 and 13 has a 3 cm ratio; links with channels within this circle are considered short or homologous connection.

For each type of connection we selected the 30 strongest in terms of their average correlations coefficient, and those ones were used to computed the indexes. By doing so, first, we guarantee having the same number of connections per type, and second, we avoid including correlations that are too weak and only add noise. The indexes were calculated as:

$$I_{Intra-Inter} = \frac{intra - inter}{intra + inter} \tag{3.1}$$

$$I_{Hom-nonHom} = \frac{hom - nonhom}{hom + nonhom} \tag{3.2}$$

$$I_{Right-Left} = \frac{rr - ll}{rr + ll} \tag{3.3}$$

$$I_{Short-Long} = \frac{short - long}{short + long} \tag{3.4}$$

In order to compare the different types of connections we contrasted each index computed on the strength and on the temporal variability against chance level (zero) using two tails t-test. P-values were Bonferroni corrected.

Furthermore, we investigated if the relative strength or variability of the different type of connections predicted the differential response towards Words and Part-words during test blocks, which can be considered as a measure of task performance To do so we computed Spearman correlations between the indexes for the strength and variability of the connections during familiarization, and the differential activation during test blocks. Because we wanted the differential activation to be as robust as possible, we defined it in terms of HbO_2 that presents a bigger signal to noise ratio, and during the time in which the cluster based permutation analysis showed consistent differences between conditions. We considered the period $[+10s, +30s]$ respect to the onset, and given that the response was broad, across all the channels. We defined the differential activation as:

$$DiffAct = \sum_{i}^{ch} \frac{Act_{i,w,HbO_2} - Act_{i,p,HbO_2}}{|Act_{i,w,HbO_2}| + |Act_{i,p,HbO_2}|} \tag{3.5}$$

where Act_{i,w,HbO_2} is the activity towards words in channel i, and Act_{i,p,HbO_2} the activity for part-words in channel i.

Principal Components Analysis. We used PCA in order to identify common patterns of variation of the correlation coefficients between pairs of channels across time and across subjects. PCA is a statistical method that enables the identification of the principal dimensions in which the data varies by performing a rotation of the coordinate system. The variation of n variables is expressed in terms of n orthonormal vectors (eigenvectors or components), obtained by linear combination of the original variables. The original variables are combined in such a way that the first component explains the maximum amount of variance, the second component explains the maximum of the still unexplained variance, and so forth. If D is a matrix of size $n \times m$ containing m observations of n variables, to apply PCA means to find W and Λ, such that $DD^T = W\Lambda W^T$, where W contains the weights to build the eigenvectors in its columns and Λ is a matrix of $n \times m$ containing the eigenvalues or scores associated to each component.

We applied PCA to the time course of the correlation coefficients for all subjects. To do so we implemented a procedure used in a previous work by Leonardi et al., 2013, which aims to identify common patterns of variation across connection pairs beyond difference between subjects. The steps were the following (see Figure 3.3): 1. We took the upper part of each Fisher transformed correlation matrices (the matrices are symmetric) and we obtained a vector of length 276, $C_{s,t}$, per subject per time window. 2. For each subject, we concatenated the vectors

along the time dimension, $S_s = [C_{s,1}, C_{s,2}, \dots C_{s,t}, \dots]$ (276 correlation coefficients observed in 96 time windows for subject s). 3. We z-scored the data per subject. 4. We subtracted the row means, $S'_i = S_1 - \bar{S}_i$. 5. We concatenated together along the time dimension the data from all the subjects, obtaining a matrix $D = [S'_1, S'_2, \dots S'_n]$ to which PCA was applied.

D contains the change of each correlation coefficient for all the subjects respect to its mean, or in other words the dynamic of the connectivity. By applying PCA to D we obtain 276 eigenvectors (or components) and its corresponding eigenvalues (or scores), such as the dynamic of the correlation coefficients can be express as $S_s = W\Lambda_s$, where Λ_s is a matrix of size 276×96 containing the scores $\lambda_{s,v,t}$ for component v at the time window t.

So far we had only re-written the data in a smarter way, such as if some of the variables are linearly correlated, the first components will explain an important part of the variation of the data set. If so, we can approximate the dynamics by the first k components, which implies a reduction in dimensionality, and the identification of correlation pairs that vary together. The choice of the number of components to considered is arbitrary but it is affected by several factors. A bigger number of components enables to explain more variability, but at the same time it makes more likely to include components associated with noise or with patterns of variation that are not shared across subjects. Moreover the number of components to consider is limited by the further statistical analysis. As closer the number of dependent variables gets to the number of observations (subjects) it becomes more difficult to find significant results.

Functional networks topology: One of the goals of the analysis is to identify and characterize functional networks. The first eigenvectors (components) are obtained by linear combination of the original variables, and the weights of the linear combination are the contributions of each correlation coefficients to an eigenvector or activation mode. We used these weights to build functional networks associated to each eigenvector. Hereafter, I will call these networks *eigen-networks*. By applying graph theory to these networks, we can do a formal description of the topologies of the activation modes and compare them. It is worth noting that in order to compare graph measures from different networks, the networks should have the same size and mean degree (number of nodes and edges), otherwise the graph measures should be normalized by values estimated for a null hypothesis (Wijk, Stam, and Daffertshofer, 2010). To avoid having this issue, we decided to considered a fixed number of links for each eigen-network to ensure a constant mean degree. Considering those links we built unweighted undirected networks. We first reshaped the weights for each eigenvector into symmetric 24 x 24 matrices. Afterwards we thresholded and binarized the matrices by taking the top 20% of positive and the 20% of negative weights, and we used them to obtain the networks. It is important to remark that these two networks are actually part of the

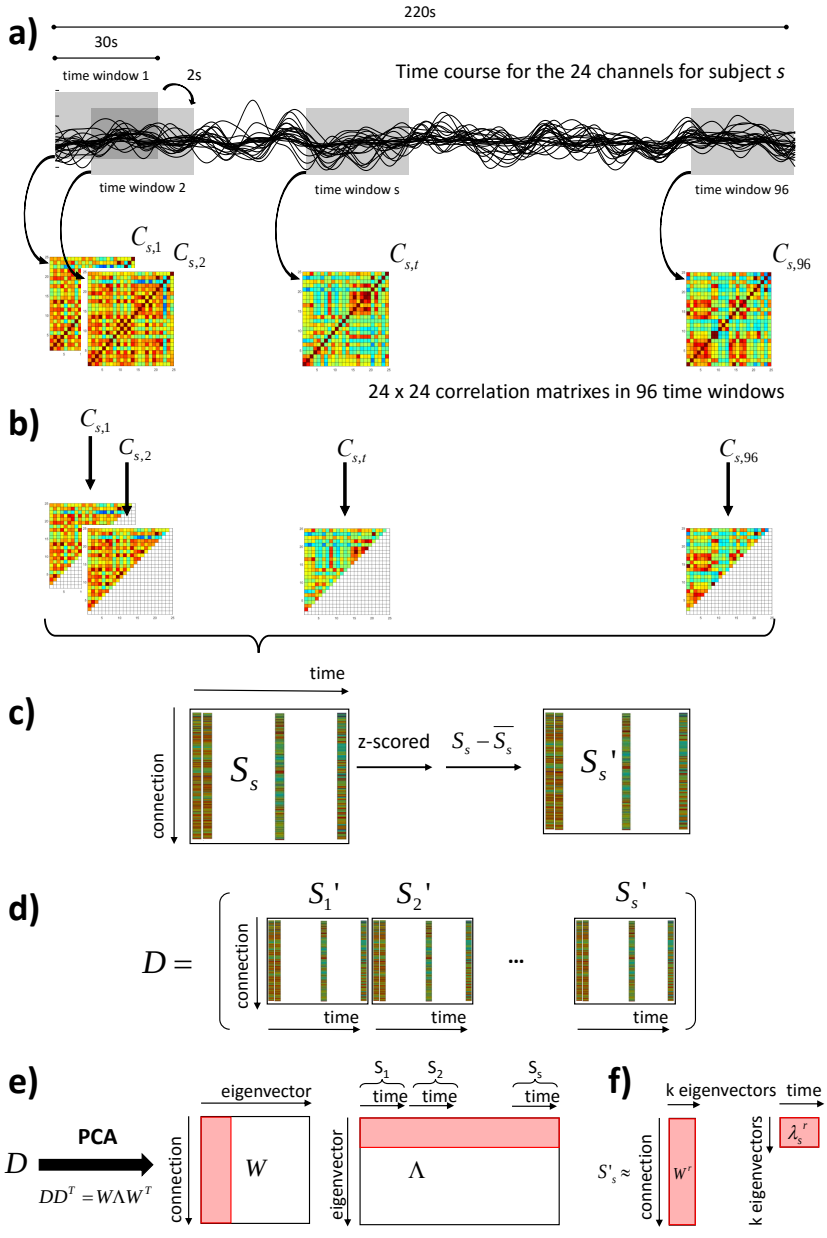


FIGURE 3.3: Pipeline of the PCA functional connectivity analysis. **a)** Temporal series for the activity in each of the 24 channels for one subject. Pearson correlations were calculated in 30s length sliding time windows. **b)** The correlation matrices were vectorized considering their upper part. **c)** The vectors from all time windows were concatenated obtaining a matrix of size 276 connections x 96 time windows for each subject and the row mean was subtracted. **d)** The matrices of the different subjects were concatenated along the time dimension obtaining a matrix, D, for the whole dynamic. **e)** PCA was applied on D. **f)** The dynamic for each subject was approximated by the first k eigenvectors. The corresponding eigenvalues for each subject were separated to describe the dynamic per subject.

same activation mode, but comprise correlations that increase or decrease their magnitude in opposition: when one increases the other decreases and vice versa depending of the sign of the eigenvalue. Moreover, the sign of the weights (and of the networks) and eigenvalues are arbitrary and can be reverted. We finally described the topology of the eigen-network using Graph Theory measures. The measures we chose are:

Ck betweenness centrality to quantify the existence of central nodes connecting different hubs.

C mean clustering coefficient as a measure of functional segregation.

E global efficiency as a measure of functional integration

sw small world to evaluate the compromise between segregation and integration.

We used the Brain Connectivity Toolbox (http://www.brain-connectivity-toolbox.net/) to estimate the measures. We are dealing with unweighted undirected networks, hence we used the following functions: *betweenness_bin* to calculate the betweenness centrality of each node; *clustering_coef_bu* to obtain the cluster coefficient of each node, the mean cluster coefficient is the average across nodes; *distance_bin* to calculate the distance between nodes and base on it *charpath* to estimate the global efficiency and the characteristic path length (L). L was used to obtain the small world as $sw = C/L$.

We should note that to build the networks we considered a node a channel. Even if we cannot be sure that each channel represents an area with a coherent activity and that there is not overlap between them, the set of nodes were always the same, thus we can still compare the network.

We also performed the analysis using other thresholds (10% and 30% stronger weights), without substantial changes in the results.

Dynamics of the activation: A second goal of the analysis is to describe the activity of functional networks; and ultimately to find differences in the dynamic of the networks that will predict the differential activation for Words and Part-words during test blocks. The dynamic of the activation for each subject is represented by the eigenvalues ($\Lambda_{v,s,t}$ is the contribution of the eigen-network v for subject s at time t) hence, differences should be reflected in differences in the distribution of the eigenvalues across eigen-networks and subjects. At this point it is important to notice that the weights we used to build the eigen-networks can be positive or negative. In this context the eigen-networks can be seen as modes of activation, with sets of connections showing opposite behaviours (connections with positive vs. connections with negative weights). The direction of the behaviour at any time is determined by the eigenvalue: a positive eigenvalue implies an increase in connectivity of the positive part of the eigen-network and a decrease of the negative part; whereas a negative eigenvalue implies the opposite.

We used two indexes to described the dynamic:

P explained variability. It is the normalized sum of squares of the eigenvalues, and it represents the total activity of the eigen-network.

$$P_{v,s} = \frac{\sum_{t=1}^{T} \Lambda_{v,s,t}^2}{\sum_{v=1}^{276} \sum_{t=1}^{T} \Lambda_{v,s,t}^2} \tag{3.6}$$

A asymmetry (or skewness). It is the difference between the mean and the median of the eigenvalues. It is positive if the distribution of the eigenvalues has a right tail, and negative if it has a left tail. It represent a bias for an activation of the eigen-network in one direction. For example a positive skewness means that the positive part of the eigen-network reaches moments of higher activity than the negative.

$$A_{v,s} = \frac{\overline{\Lambda_{v,s}} - \widetilde{\Lambda_{v,s}}}{std(\Lambda_{v,s})} \tag{3.7}$$

For the sake of investigating a relation between functional connectivity during the familiarization phase and the later activity toward Words and Part-words, we calculated the Spearman correlations between P and A and the differential activation for Words and Part-words during test blocks.

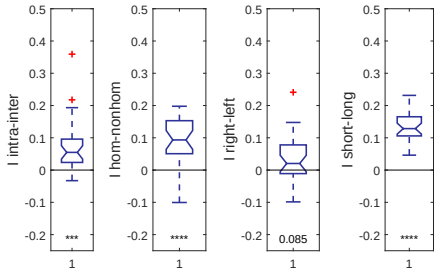
Phase randomization simulations: The sliding time window method can result in spurious functional connectivity dynamics (Handwerker et al., 2012), thus it is necessary to compare results with a null hypotheses to verify that the observed results actually depend on the precise timing between time series. To do so we run phase randomization simulations. We took the pre-process time course of the brain activity and we phase randomized it independently. In more detail, we Fourier transformed the time series, added a random phase independently to each time series, and transformed back to the time domain. Afterwards, we band-pass filter the generated data set, and we performed the connectivity analysis as described before. We repeated the operation 100 times. The simulated data set was our null hypothesis. Because the synthetic time series preserve the amplitude spectra and autocorrelation properties, but lose the precise timing, if the observed correlation fluctuations differ from the null hypothesis, it means they are not a product of random temporal activity.

Static functional connectivity. The aim of this final control is to evaluate the aspects of the functional connectivity that emerge when stationarity is assumed, thus by looking to the average stronger functional connections. We calculated the temporal correlation during the entire familiarization and to build the networks we binarized the correlation matrix considering the 20% higher correlations.

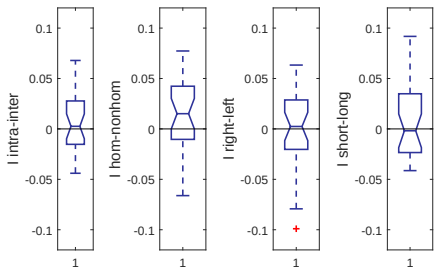
3.2.6 Results

The results for HbO_2 and Hb are presented below. Figures and tables of the Hb results are in the Appendix B.

Strength and variability of the connections. The indexes on the strength and temporal variability defined in the methods are presented in Figure 3.4 for HbO_2 and Figure B.2 for Hb. We first investigated if there were difference between connections for the infants as a group by comparing the indexes against chance level (zero). By using HbO_2 to estimate functional connectivity, two tails t-tests against zero revealed that: intra hemispheric connections were stronger than inter-hemispheric (P_{corr} < 0.0001); that homologous connections were stronger than non-homologous (P_{corr} < 0.0001), and that short intra-hemispheric connections were stronger than long intra-hemispheric connections (P_{corr} < 0.0001). Right intra-hemispheric connections were marginally significant stronger than left (P_{corr} = 0.0850) (see Table 3.1). Using Hb we found a similar pattern of results but no difference was found between the strength for right and left connections (see Table B.1). Regarding the variability, the only significant difference was between intra-hemispheric connections and inter-hemispheric connection in terms of Hb (P_{corr} < 0.05), meaning that intra-hemispheric connections were more variable along time.



(A) Strength



(B) Variability

FIGURE 3.4: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 1. Connectivity wasro. estimated using HbO_2. Asterisks bellow represent Bonferroni corrected P-values of the t-tests against ze

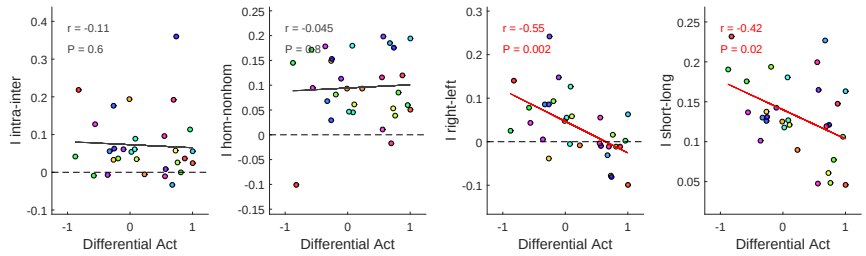
We also investigated if the differential activation for part-words and words during test blocks could be predicted by some connections being stronger or more

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength HbO_2	I intra-inter	0.0716	0.0846	0.0400	0.1031	29	4.6343	0.0001	0.0000
	I hom-nonhom	0.0956	0.0702	0.0694	0.1218	29	7.4629	0.0000	0.0000
	I rigth-left	0.0327	0.0735	0.0052	0.0601	29	2.4344	0.0213	0.0850
	I short-long	0.1329	0.0495	0.1144	0.1514	29	14.7072	0.0000	0.0000
Variability HbO_2	I intra-inter	0.0067	0.0300	-0.0045	0.0179	29	1.2191	0.2326	0.9300
	I hom-nonhom	0.0088	0.0351	-0.0043	0.0219	29	1.3756	0.1795	0.7180
	I rigth-left	0.0009	0.0404	-0.0142	0.0160	29	0.1200	0.9053	3.6210
	I short-long	0.0070	0.0349	-0.0060	0.0201	29	1.1066	0.2776	1.1100

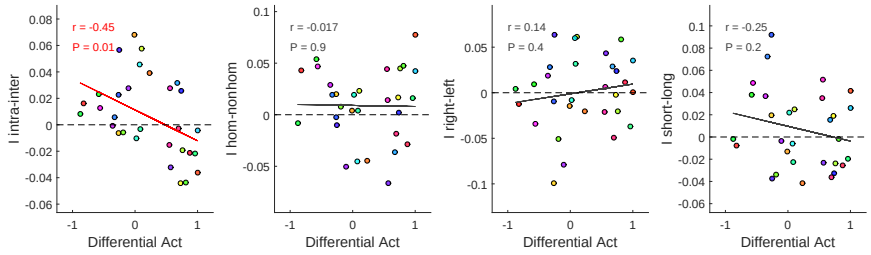
TABLE 3.1: Statistical analysis on the different indexes For Experiment 1 using HbO_2. T-tests against chance (zero). Bonferroni correction was used.

variable across time respect to others. To do so we correlated the asymmetry indexes with the differential activation during test blocks. Spearman correlations showed a negative correlation with the righ - left asymmetry index and (R = 0.5509, P < 0.01) and with the short - long index (R = -0.4250, P < 0.05). In terms of the temporal variability of the connections, the differential activation was negatively correlated with the intra-inter asymmetry index (R = -0.4494, P < 0.05) (see Figures 3.5). All the other correlations were not significant (P > 0.05). Using Hb we observed similar trends but results were noisier. We found a significant positive correlation with the strength of the homologous - non-homologous index (R = 0.4178, P < 0.05). In brief, we can say that infants that show a bigger differential activity for part-words and words, also presented stronger left-left functional connections respect to right-right connections, stronger long than short intra-hemispheric connections, and higher variability in inter-hemispheric connections across time (or lower variability in intra-hemispheric connections).

Principal Components Analysis. Applying PCA we found 276 new orthonormal variables (eigenvectors) to represent the data, obtained in such a way that each of them explains the maximum of the remaining variance of the data. In Figure 3.6.a and Figure B.4.a the explained variance is plotted against the eigenvector number together with the data from the phase randomization simulations. We would like to focus the analysis on the relevant eigenvectors, meaning the ones representing real common patterns of variation. To do so, we compared the variance explained by the eigenvectors obtained from the real data, with the null distribution obtained from the phase randomization simulations. With a significance level of 5% we estimated that only the first 20 eigenvectors for HbO_2 and 16 for Hb, explained more variance of what we could expect if no temporal structure was present in the data. The descriptive graph measures for the real data for the first components were also higher that the ones obtained from the phase randomized data (see Figure 3.6.b-e and Figure B.4.b-e), whereas for the rest components they fell inside the confident intervals given by the simulations. This provides further evidence that the first components capture common patterns in the functional dynamics that are not attributable to random noise. Moreover, it shows that



(A) Strength



(B) Variability

FIGURE 3.5: Spearman correlations between task performance (differential activation during test blocks) and functional connectivity for Experiment 1 using HbO_2. The differential activation for part-words and words during test blocks (x-axis) is plotted against the different indexes estimated during the familiarization phase (y-axis): $I_{inter-intra}$, $I_{hom-nonhom}$, $I_{right-left}$, $I_{short-long}$. Each dot represent a subject.

.

its associated networks present an efficient organization.

In order to have enough statistical power we needed to reduce our analysis to a smaller set of variables, therefore we decided to focus on the first four eigenvectors. For these eigenvectors all the descriptive graph measures used are above the 95 percentile obtained from the simulated data (see Figure 3.6.b-e). Together the first four eigenvectors explained 31.0% of the variation of the data for HbO_2 (eigenvectors 1 to 4 explained respectively 22.1%, 3.3%, 3.0% and 2.6%) and 37.6% for Hb (eigenvectors 1 to 4 explained respectively 27.9%, 3.6%, 3.2% and 2.9%). These values differ drastically by the amount of variance explained by the first 4 eigen-networks in the phase randomization simulations, 5.69% for HbO_2 (1.50%, 1.44%, 1.39% and 1.36% respectively) and 5.78% for Hb (1.52%, 1.46%, 1.42% and 1.38% respectively). The next 4 eigen-networks (5 to 8) are presented in Appendix B (see Figure B.1) with a illustrative purpose.

Functional networks topology: We thresholded and binarized the weights of each eigenvector to build unweighted functional networks, or eigen-networks. In order to characterize the topological organization of the networks we calculated for each node its degree (number of connections), cluster coefficient (proportions of triangles) and centrality (shortest paths that pass through that node). In addition

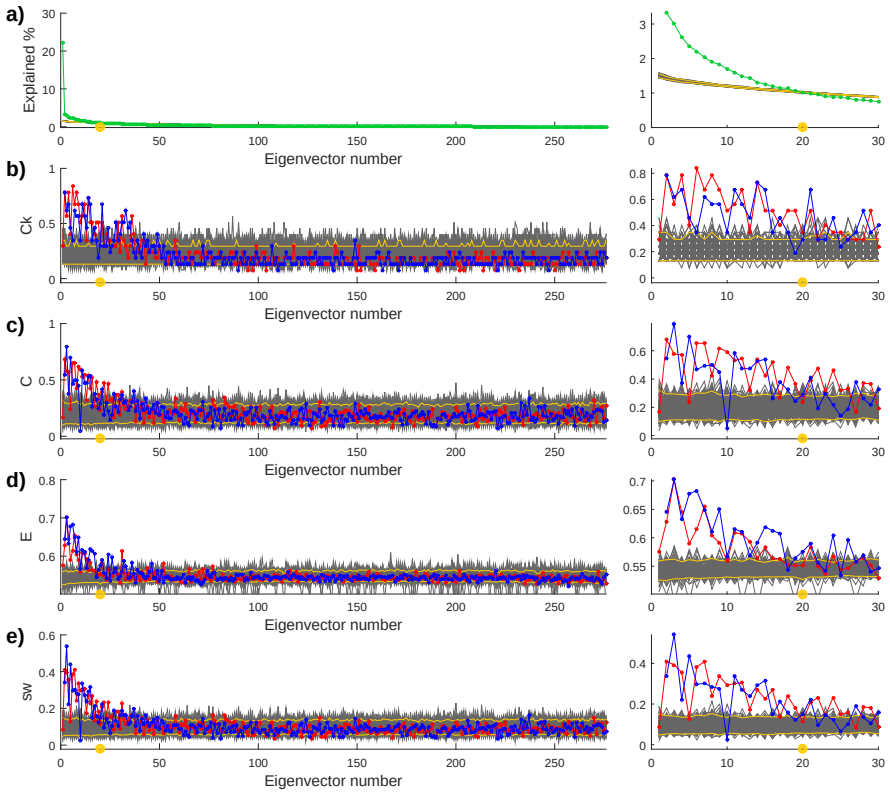


FIGURE 3.6: Phase randomization simulations for Experiment 1 using HbO_2. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

we used four descriptive graph measures for the whole network: betweeness centrality, mean cluster coefficient, global efficiency and small-world. In Figure 3.7 the matrices with the weights (a), the eigen-networks using the 20% top weights (b), and the z-score graph measures for each eigen-network (c) are presented.

The first eigen-network has only positive weights and represents global changes in the functional connectivity. In accordance with this, eigen-network 1 explains most of the variance in data. The rest of the eigen-networks have positive and negative weights, meaning that are activation modes with functional connections showing opposite behaviours —when the synchronization between some pairs of channels increases the synchronization between others consistently decreases. In eigen-network 2, parietal inter-hemispheric connections behave oppositely to connections involving more fronto-temporal regions, with nodes of high degree in left temporal areas. Eigen-network 3 comprises connections involving more

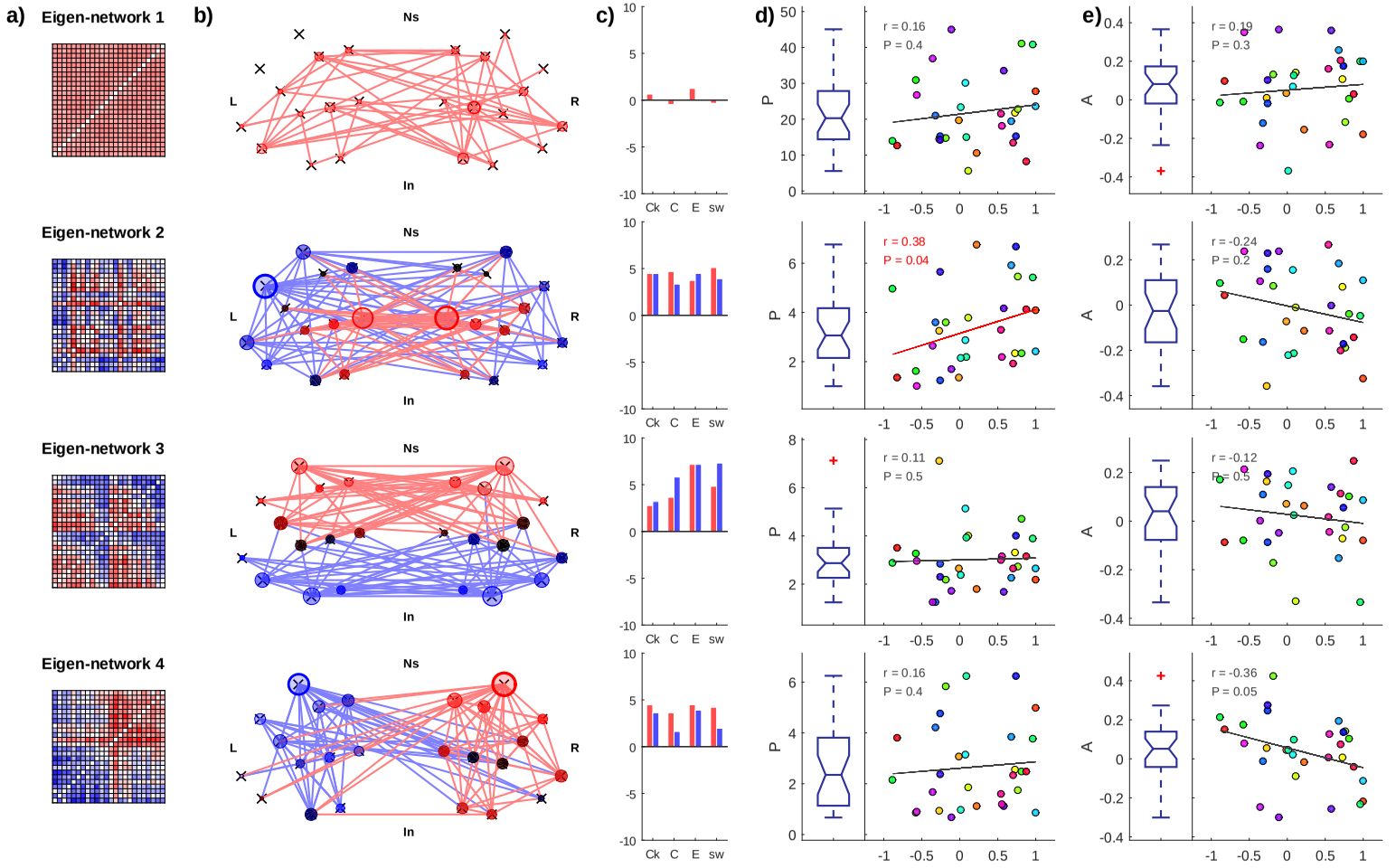


FIGURE 3.7: PCA analysis for Experiment 1 using HbO_2. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z-scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Spearman correlations between the explained variability (P) for each eigen-network and the differential activation for Part-words and Words. **e)** Spearman correlations between the asymmetry (A), for each eigen-network and the differential activation for Part-words and Words.

frontal areas in opposition to connections between more posterior regions. Finally, eigen-networks 4 comprises two remarkably symmetric networks, one involving mainly right intra-hemispheric connections and a frontal right node connected with left temporal areas, and a mirroring network for the left hemisphere.

Regarding the measures describing the whole network topology (centrality, mean cluster coefficient, efficiency and small-word), we can observe that eigen-networks 2-4 present values substantially above eigen-networks explaining less variance (z-score values of approximately 2 to 5), and also significantly above than the networks from the phase randomization simulations (see Figure 3.6). These results reveal that the topological organization is not consistent with random noise networks, but that a particular architecture exists. Meanwhile, eigen-netwok 1 presents graph measures that are around the mean, indicating it does not have a particular organization. This agrees with the hypothesis that it represents global changes in the functional connectivity —its weights are all similar— and with Leonardi et al., 2013 findings.

When we inspected the architecture of the last eigen-networks (see Figure 3.8), we observed that they lack of organization. Unlikely eigen-networks 2-4 that show a well-organized architecture and are likely to represent real fluctuation in connectivity, the last eigen networks can probably be attribute to noise. In fact they structure is as messy as for the first eigen-networks from the phase randomization simulations (see Figure 3.9).

The same pattern of results was observed considering the 10% and 30% top weights to build the eigen-networks, confirming that the results do not depend of the chosen threshold (see Figures 3.10).

Using Hb we found eigen-networks that resemble the ones found for HbO_2, but again results were noisier (see Figure B.5).

Dynamics of the activation: The dynamics of the functional connectivity is described by the eigenvalues. The eigenvalues can be interpreted as the activation of each eigen-network at each moment. For example in Figure 3.11 the time course of the eigenvalues for the first four eigen-networks are shown for one subject, together with the 5th and 95th percentiles estimated from the phase randomized simulations. The eigenvalues oscillate below and above the 5th and 95th percentiles for all subjects indicating: 1. that the functional connectivity highly fluctuates along time; 2. that the eigen-networks are representative pattern of fluctuations common to all subjects.

Eigen-network 1 may represents global changes in connectivity, and in fact this was confirmed by its eigenvalues being strongly correlated with the mean correlation over all connections pairs along time (R = 0.999, P < 1e-110 for all subjects) —when connectivity is strong the eigenvalues for eigen-network 1 are highly positive, whereas when connectivity is weak they are strongly negative.

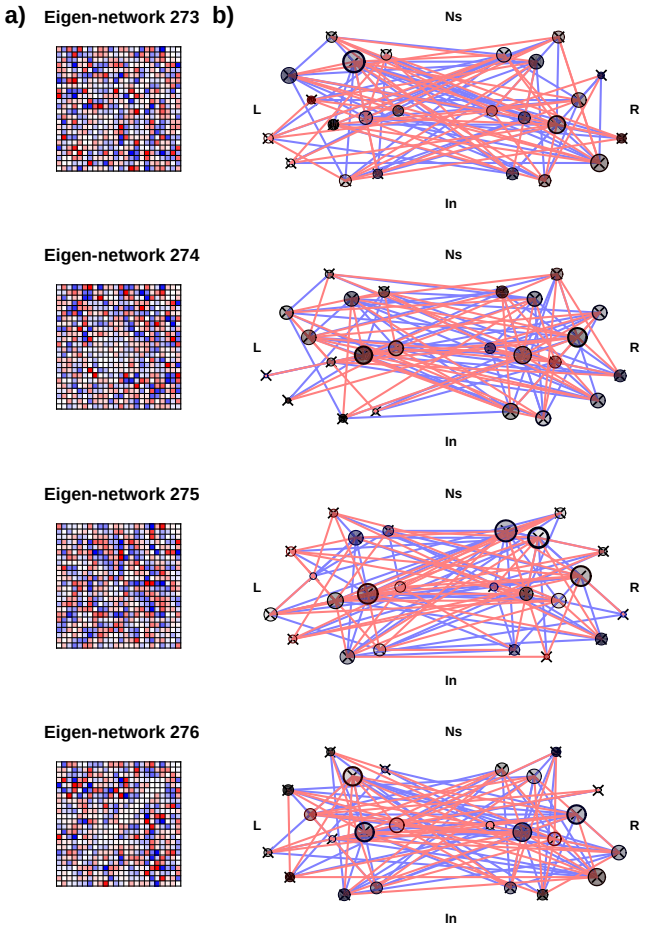


FIGURE 3.8: Last four eigen-networks of the PCA analysis for Experiment 1 using HbO_2. The representation uses the same code than in Figure 3.7.

Going further, we wanted to see if the dynamic of the activation differed between good and bad performers. As I said before, we used two measures to characterize the eigenvalues' distribution for each of the first four eigen-networks: the percentage of the variability explained (P), and the asymmetry of the distribution (A). In panel (d) and (e) of Figure 3.7 the Spearman correlations between these measures and the differential activation during test blocks are presented. We did not find strong effects, but we found some marginally significant results in line with the rest of the results, and in particular consistent with the correlations between performance and the indexes defined for the mean activation analysis. The differential activation for part-words and words was significantly positively correlated with the total activity of eigen-network 2 (R = 0.3820, P = 0.0372), and marginally negatively with the asymmetry for eigen-network 4 (R = -0.35828, P = 0.0519). Results were not significant when corrections for multiple comparisons was applied.

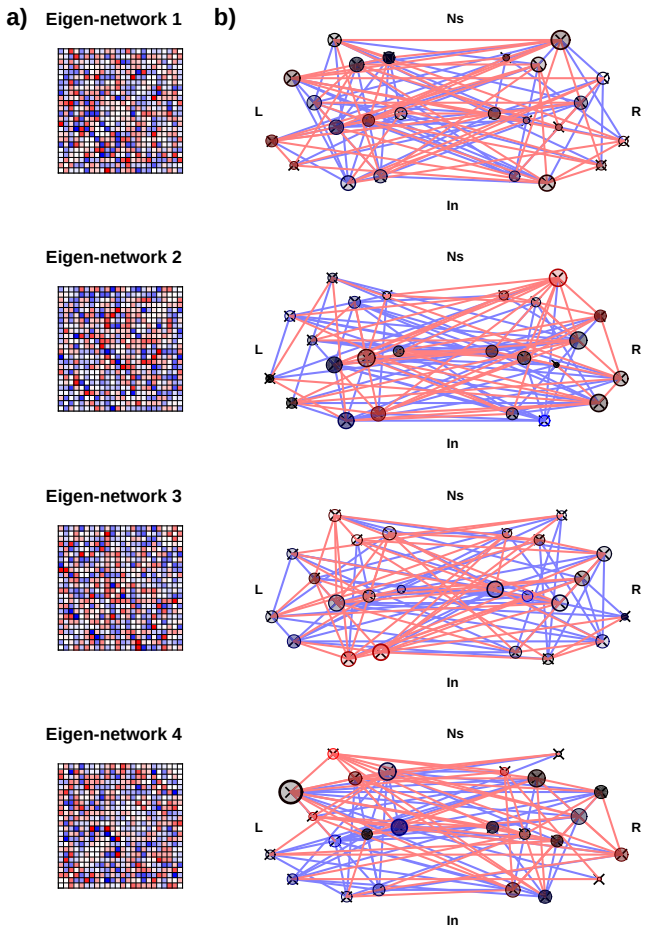


FIGURE 3.9: First four eigen-networks from one phase randomization simulation for Experiment 1 using HbO_2. The representation uses the same code than in Figure 3.7.

Static functional connectivity. We performed this analysis in order to compare results between an analysis exploring the temporal aspect of the functional connectivity, and results obtained by assuming stationarity. The Fisher transformed correlation matrix for the entire familiarization and the functional network build using the top 20% correlations are presented in Figure 3.12. Observe that the functional network obtained resembles eigen-network 1.

3.2.7 Discussions

The main aims of our analysis were: 1. To describe the organization of the cortical activity over the recorded areas during the task. 2. To explore its dynamics and describe functional networks based on it. 3. To find aspects of the functional connectivity that correlate with task performance.

To do so we implemented two analysis based on a sliding time windows approach. The first analysis directly described the strength and variability of the

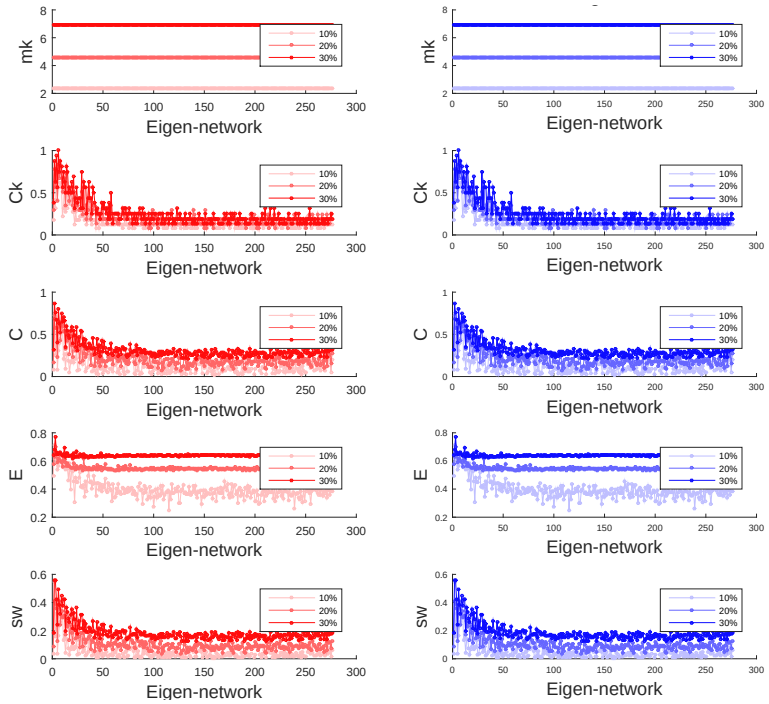


FIGURE 3.10: Graph measures as function of the eigen-network number for Experiment 1 using HbO_2. The measures for the positive part of the eigen-netwok are shown on the left (red) and for the negative on the right (blue). The results using different thresholds (10%, 20% and 30 % top weights) are shown in different colours. Mean degree (mk). Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw).

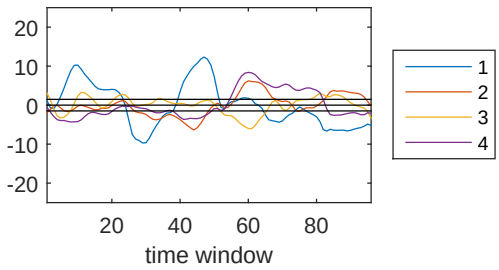


FIGURE 3.11: Eigenvalues along time for one subject for the first four components for Experiment 1. Results for HbO_2 are shown.

functional associations; whereas the focus of the second analysis was on the identification of functional networks based on the variations in connectivity.

The first analysis revealed that functional connectivity was stronger for intra-hemispheric than inter-hemispheric connections, and that short range intra-hemispheric connections were stronger than long range connections, which agrees with finding suggesting a more local organization at birth (Fransson et al., 2011; Perani et al., 2011). Among hemispheres, functional connections between homologous

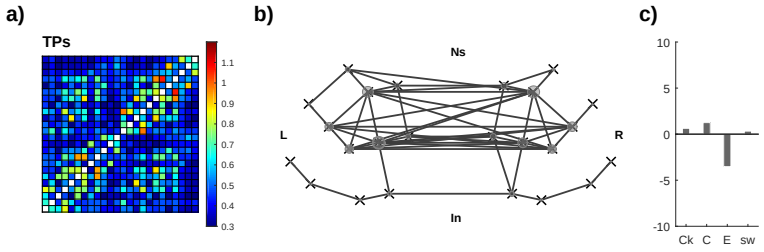


FIGURE 3.12: Static functional connectivity analysis for Experiment 1 using HbO_2. **a)** Fisher transformed correlation coefficient matrix. **b)** Network obtained by thresholding (top 20 %) and binarizing the correlation matrix. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z-scored graph measures for the network. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw).

regions were clearly stronger, another phenomenon previously observed in connectivity based on BOLD signal in infants (Homae et al., 2010). These two results may have its origin in the earlier development of tracks connecting hemispheres than anterior posterior fibres (Dubois et al., 2008; Dubois et al., 2014; Perani et al., 2011). For right-left asymmetry even if we did observe a trend in the direction of stronger right connectivity —something compatible with an earlier maturation of the right STS during the first months of life (Leroy et al., 2011)— the difference was not significant. In terms of the temporal variability of the connections, we did not find differences. Notice that this does not mean that the connections we compared change together, but simply that present a similar level of variation along the task.

Meanwhile, the PCA analysis enabled us to identify the common patterns of variation in the fluctuation of the functional connectivity. We found out that the main source of variability was due to global fluctuations in connectivity, which was reflected in the huge amount of variance explained by eigen-network 1. We also identified variation modes —eingen-networks— that explained more variability of the data set of what would be expected from random noise. These modes were conformed by groups of connections with opposite behaviours, reflecting that more synchronized activity between some areas is in general associated with a decrease in synchronization between other areas. In particular we found an opposition between: more frontal and more posterior regions, left and right hemispheres, and between parietal and more fronto-temporal areas. A remarkable aspect is that these eigen-networks present small-world topological properties that were not present in eigen-networks associated to noise —either eigen-networks explaining less variance and networks obtained from the phase randomized data. In brief, the networks extracted from the PCA analysis revealed a functional organization dynamic with a very specific structure since birth.

Our final goal was to identify aspects in functional connectivity that predict task performance. We did not have a direct measure of task performance, but we assumed that infants that segmented and extracted better the words should show bigger differential response for words and part-words during test blocks. We used this differential activation as measure of performance —note that this measure is independent from our connectivity measures. Based on our first analysis we found that infants with stronger left intra-hemispheric connections and stronger long range connections were better at the task. These results go in line with previous literature signalling a dominance of the left hemisphere for speech perception since birth (Peña et al., 2003; Mahmoudzadeh et al., 2013), and furthermore with the presumable need of integrating information in different areas for a task of this nature. Previous work shows that whereas the STG is already connected with the IFG though the ventral pathway and to the motor cortex through a dorsal path; the dorsal pathway connecting the temporal area and Broca's area is not developed at birth yet (Perani et al., 2011). If the segmentation task requires integrating information between the STG and the IFG, our results may reflect that infants performing better have a higher developmental of these fibres, providing a functional phonological loop to generate short term representations of the words, which shows up in the presence of stronger long range intra-hemispheric connections. We also observed bigger temporal variance for intra-hemispheric connections relative to intra-hemispheric in bad performers. The origin of this effect is not clear, but may be due either to more stable intra hemispheric connections in good performers, or to a further transient increases in the strength of inter-hemispheric connection, which would lead to more variation.

In the PCA analysis, the eigen-networks are the same for all the subjects, hence the prediction of performance should be based on the dynamic of the activation of the different eigen-networks. Even if we did not find strong results, we observed some trends consistent with the results barely described. On one hand, infants with a better performance showed stronger activity for eigen-network 2, which involves many inter-hemispheric connections either between parietal areas (positive part) or more temporal regions (negative part). On the other hand, for good performers, eigen-network 4 that dissociated right (positive part) and left (negative part) hemisphere, showed an asymmetry towards higher activity of the left hemisphere. The lack of strong prediction of performance by the eigen-networks activity could be consequence of none of them specifically and fully involved in the task. As it will be discussed for the next experiments they may reflect functional networks that are not task dependent.

All together, the correlations between functional connectivity and task performance where consistent across analysis and with previous anatomical studies, suggesting a real neural bases of the results.

A final aspect to remark is the similarity between the functional network obtained from correlations during the whole familiarization —assuming stationarity— and eigen-network 1. By computing correlations along the whole period we obtained an estimation of the average connectivity, whereas eigen-network 1 shows that global and strong fluctuations in connectivity exist. The fact that the two networks resemble suggests that the global variation entails changes for each connection that are proportional to their strength. The interesting aspect of the PCA method is that regardless the global fluctuations, it enables to identify coordinate pattern of variation that are missed by a stationary connectivity analysis. Assuming stationarity means losing the temporal dimension and therefore the possibility of identifying synchronized activity besides the areas with globally strong connectivity. The average level of synchronization may be affected by distinct factor; for example two areas could have similar function but their activity present a fix phase delay, leading to a lower correlation. The method is a second order approach and senses coordinated changes from a “baseline" connectivity level.

Summarizing, the functional connectivity analysis on Experiment 1 revealed: 1. Strong functional associations between homologous regions and between spatially closed areas. 2. Strong variations in the global brain connectivity. 3. Functional networks with small-world properties and opposite activation patterns. 4. A correlation between functional connectivity and task performance —better performers had stronger synchronization between areas of the left hemisphere and between distant regions within hemisphere, and likely more stable intra-hemispheric connections.

These findings open many questions. First, we cannot say which activation is task related and what is common with resting state activity. Second, and related with the previous, the correlation between task performance and some features of the functional connectivity may be due to individual differences in brain architecture —probably as consequences of anatomical and functional differences in development— and hence observable also during rest; or they may have been induced by the task. In order to to have a more complete picture we conducted a control experiment.

3.3 Comparing functional connectivity while listening structure and random sequences of syllables and rest. Experiment 3

With this control experiment we tried to understand which aspects of the dynamics of the functional connectivity were associated with the segmentation task. To do so a new group of neonates were exposed to three conditions each of them lasting the same duration as the long familiarization of Experiment 1. Infants heard a period of silence (Sil condition), a sequence of random syllables (Rnd condition)

and a structured sequences of syllables as the one of Experiment 1 (TPs condition). The Rnd condition is auditory similar to the TPs condition but lacks any structure, hence it enable us to test whether the previous results were due to the structure in the stream or to hearing a sequence of syllables. Importantly the set of syllables used to build the Rnd and TPs stream were different.

We applied the same two analysis than for Experiment 1. In the first analysis we directly compared the strength and variability of the connections across conditions. The PCA analysis was applied in the exact same way than before, but across the three conditions together, meaning that we applied PCA to a connectivity dynamic obtained from all subjects during all the conditions, thus we obtained a single set of eigen-networks. The differences between conditions should be reflected in the activity of the eigen-network.

We hypothesised that if the functional networks and the differences in functional connectivity observed in Experiment 1 are task related, differences across the conditions of Experiment 3 should be observed. In particular, aspect of the dynamic functional connectivity related to computing the transitional probabilities and extracting the words should be reflected in differences between the TPs and Rnd conditions; whereas, functional connectivity changes due to hearing a sequence of syllables should appear between the silence period and the other two conditions. If it is the case that our measures reflect functional associations that are task-independent —either due to direct anatomical connections or pure functional associations— no differences should appear across condition. In particular we expect differences in the strength and variability of the connection for the first analysis and in the distribution of the eigenvalues —reflected in P and A— for the PCA analysis.

3.3.1 Participants

All participants were healthy full-term neonates born to Italian-speaking mothers, with Apgar score ≥ 7 in the first minute and ≥ 8 in the fifth minute, diameter of head ≥ 33.0 cm, and no cefalhematomas. Experiment 3 included 23 participants (10 females; mean age 3.04 days; range 2-5 days; mean gestational age 39.1 weeks, rage 37-41 weeks; mean weight 3.405 Kg, SD 0.421 Kg). Additional infants were tested but excluded from the final analyses because too many motion artefacts (n = 17), failure to complete the experiment due to fussiness (n = 6), or a poor signal due to thick hair (n = 6). All newborns were recruited from the nursery at Hospital, Azienda Ospedaliera Santa Maria della Misericordia, in Udine, Italy. Parents provided informed consent. The Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

Stream	TPs: words	Random: syllables
1A	lamipe duvoka nubefi telugo	pi ko vu de ma ti fe bu lo gu le na
1B	mipedu vokanu befite lugola	pi ko vu de ma ti fe bu lo gu le na
2A	pivoku demati febulo gulena	la mi pe du vo ka nu be fi te lu go
2B	kovude matife bulogu lenapi	la mi pe du vo ka nu be fi te lu go

TABLE 3.2: Stimuli for Experiment 3. Four familiarization streams and sets of Words and Part-words were used, and infants were randomly assigned them.

3.3.2 Stimuli

Stimuli were synthesized using the it4 Italian female voice of the MBROLA diphone database (Dutoit et al., 1996), with phoneme duration of 150 ms and a constant pitch of 200 Hz. Sequences were continuous with no pauses between syllables.

The TPs streams had identical structure than the long familiarization stream of Experiment 1. The Random syllables streams were created concatenating semi-randomly 12 syllables, with the only restriction that the same syllable could not be repeated twice and that two syllables could not alternate more than three times (the sequence *A–B–A–B*, where *A* and *B* are two syllables, was not allowed). As result we obtained a stream with each syllable appearing with a frequency of $1/12$ and uniform transition probabilities of $\approx 1/11$. The average transition probability between syllables was 0.0908 (SD = 0.0350, range = [0.0167, 0.1833]). In order to avoid that results were driven by low level properties of the stimuli, four TPs streams and two Random streams were created using the sets of syllables of Table 3.2. Infants were randomly assigned to the groups. Notice that the set of syllables used for the TPs and Random stream were different. The syllables of the TPs stream of group 1A and 1B were used to build the Random stream of groups 2A and 2B and vice-versa.

3.3.3 Procedure

Three long blocks (TPs, Random and Silence) were presented in random order to the infants. Each block lasted 220 s and were separated by silence periods of 25-30 s. The TPs and Random block were ramped up and down for 6 s. The total duration of the experiment was 12.5 minutes (see Figure 3.13).

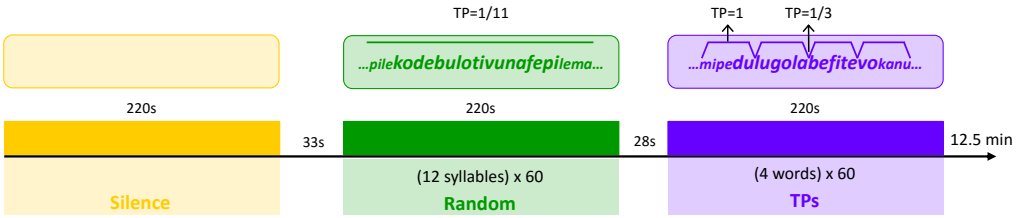


FIGURE 3.13: Schematic representation of the protocol for Experiment 3. Infants heard a period of continuous stream built concatenating in random order four three-syllabic words (TPs), a period of continuous stream conformed by twelve syllables randomly concatenated (Rnd), and a period of silence (Sil). Each block lasted for 220 s. The order of presentation was counterbalanced across subjects.

3.3.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

3.3.5 Data Analysis

The analysis was essentially the same than for Experiment 1, but in this experiment we looked for difference between the three conditions (TPs, Random, Sil).

Pre-procesing. The pre-processing was identical to the the pre-processing described for Experiment 1 (see 3.2.5).

Data rejection. We excluded subjects that presented motion artifacts during more than a 30% percent of the duration of either the TPs, Random or Silence blocks, thus included subjects contributed to all the conditions.

Stimation of the functional connectivity. Connectivity was estimated in the same way than for Experiment 1 (see 3.2.5). We calculated the Pearson correlation matrices and Fisher transform the correlation coefficients for the three conditions.

Strength and variability of the connections. We estimated the average functional connectivity as the mean of the correlation coefficients between all pairs of channels, and we evaluated the effect of condition on its temporal mean and temporal variability by a 1-way ANOVA. In addition we calculated the indexes on

the strength and variability as we did for Experiment 1 (see 3.2.5). We compared the indexes against chance level (zero) considering the data from all conditions together using a t-test, and across conditions using 1-way ANOVA.

Principal Components Analysis. We did the same analysis than for Experiment 1 (see 3.2.5). The PCA analysis was done pulling together the data from the three conditions (TPs, Rnd, Sil) for all the subjects.

Functional networks topology: We built and described the eigen-networks in identical way than for Experiment 1 (see 3.2.5). Observed that because the PCA analysis was run with the data of all the conditions, we obtained only one set of eigen-networks.

Dynamics of the activation: As for Experiment 1 we used the explained variance (P) and the asymmetry (A) to characterize the activation of the different eigen-networks under different conditions (see 3.2.5). Because the eigen-networks are the same for all the conditions, we expected to find difference across conditions in the distribution of the eigenvalues. To compare the P and A across conditions and components we fit a repeated measures ANOVA, with condition and component as within subject factors.

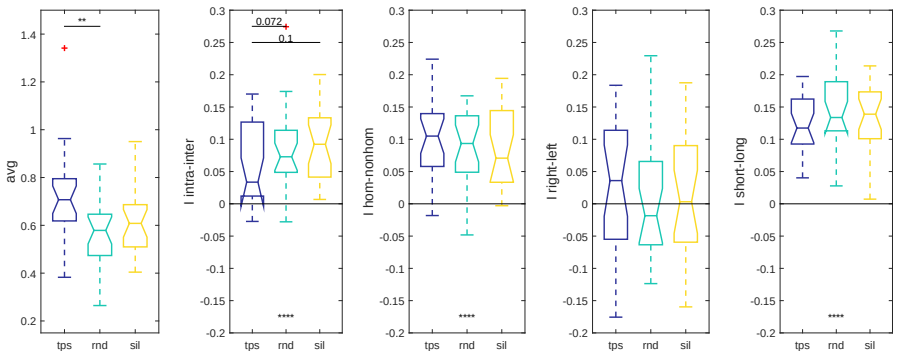
Phase randomization simulations: We performed the phase randomization simulations as described for Experiment 1 (see 3.2.5).

3.3.6 Results

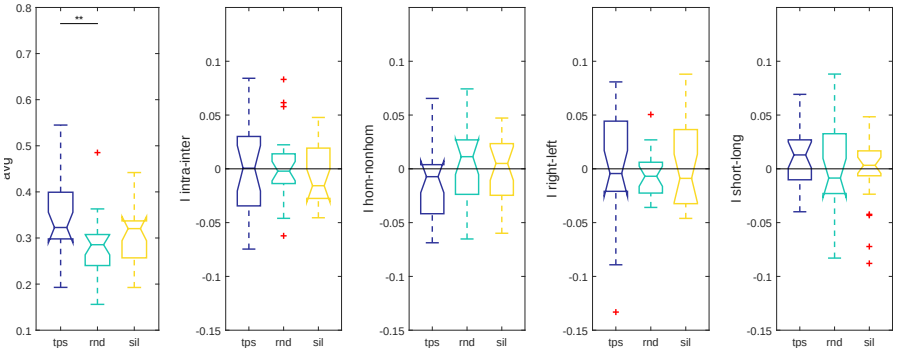
The results for HbO_2 and Hb are presented below. Figures and tables of the Hb results are in the Appendix B.

Strength and variability of the connections. First, we compared the strength and variability of the average functional connectivity, and second, we investigated for differences between the different type of connections using the asymmetry indexes (see Figure 3.14 and B.8). A 1-way ANOVA on the strength of the connections as within subject factors revealed a main effect of condition for both HbO_2 ($F(2,44) = 6.4404$, $P < 0.01$) (see Table 3.3) and Hb ($F(2,44) = 9.0554$, $P < 0.001$) (see Table B.2). Post-hoc Tukey-Kramer comparisons using HbO_2 (see Table 3.4), showed that the connectivity during the TPs condition was stronger than during the Rnd condition ($P < 0.01$), whereas no significant differences were found between the other comparisons ($P > 0.05$). For Hb, the connectivity during the TPs condition was stronger than during Sil condition ($P < 0.01$) and marginally significant stronger than during Rnd ($P=0.096$) (see Table B.3). For the variability of the mean connectivity we also found an effect of condition for both HbO_2 ($F(2,44) = 5.73406$, $P < 0.01$) (see Table 3.3) and Hb ($F(2,44) = 4.31324$, $P < 0.01$)

(see Table B.2). Post-hoc Tukey-Kramer comparisons revealed that for HbO_2 the difference was driven by higher variability in the TPs condition that was significantly higher than in the Rnd condition (P < 0.01) and marginally significant than in the Sil (P = 0.1); whereas for Hb the variability in the TPs condition was higher than during Sil (P < 0.01).



(A) Strength



(B) Variability

FIGURE 3.14: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 3. Connectivity was estimated using HbO_2. The results for each of the conditions are presented: TPs (blue), Rnd (green) and Sil (yellow). Asterisks bellow represent Bonferroni corrected P-values of the t-test against zero. Asterisks above show the results of the post-hoc multiple comparisons analysis using Turkey-Kramer correction.

		SumSq	DF	MeanSq	F	P
Strength HbO_2	Condition	0.281	2	0.141	6.44043	0.00350
	Error (Condition)	0.960	44	0.022		
Variability HbO_2	Condition	0.051	2	0.026	5.73406	0.00610
	Error (Condition)	0.197	44	0.004		

TABLE 3.3: Statistical analysis on the strength of the connections for Experiment 3 using HbO_2. 1-way ANOVA with condition as within subject factor.

By comparing the indexes against chance level we obtained the same results than for Experiment 1. For HbO_2, intra-hemispheric connections were stronger than inter-hemispheric (P_{corr} < 0.0001), homologous connections were stronger than non-homologous (P_{corr} < 0.0001), and short connections were stronger than

		Difference	StdErr	P	CI lower	CI upper
Strength HbO_2	TPs-Rnd	0.153	0.043	0.00455	0.046	0.260
	TPs-Sil	0.105	0.052	0.13157	-0.026	0.230
	Rnd-Sil	-0.048	0.034	0.35540	-0.135	0.040
Variability HbO_2	TPs-Rnd	0.065	0.019	0.00587	0.018	0.110
	TPs-Sil	0.045	0.021	0.10545	-0.008	0.100
	Rnd-Sil	-0.020	0.019	0.56256	-0.068	0.030

TABLE 3.4: Statistical analysis on the strength of the connections for Experiment 3 using HbO_2. Post-hoc multiple comparisons analysis between conditions. Turkey-Kramer correction was used.

long (P_{corr} < 0.0001) (see Table 3.5). For Hb we obtained the same results but right connections were also marginally stronger than left connections (P_{corr} = 0.051) (see Table B.4). For the variability we did non find significant differences for any of the indexes (P > 0.05) (see Table 3.5 and Table B.4). When we compared the different indexes across conditions we found a marginally significant modulation on the strength of intra - inter hemispheric connections for both HbO_2 (F(2,44) = 2.70821, P = 0.07780) (see Table 3.6) and Hb (F(2,44) = 2.58437, P = 0.08690) (see Table B.5). Post-hoc Turkey-Kramer corrected multiple comparison showed that for HbO_2 the modulations was due to $I_{intra-inter}$ being marginally smaller during TPs than during Rnd (P = 0.0718) and Sil (P = 0.0999); and for Hb to the $I_{intra-inter}$ being smaller in Rnd than in Sil (P = 0.0945). We also found a significant modulation of the $I_{short-log}$ based on the strength of the connections, but only for Hb (F(2,44) = 4.26046, P < 0.05) (see B.2). The effect was due to a higher index during Sil than during TPs (P = 0.0713) and during Rnd (P = 0.0648). In terms of variability we found a significant modulation of the $I_{right-left}$ for Hb (F(2,44) = 3.69288, P < 0.05), but we did not observe the same for HbO_2 (see 3.3). Post-hoc Turkey-Kramer multiple comparisons showed that the effect was due to a lower index during Sil than during TPs (P = 0.0595) and Rnd (P = 0.0865).

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength HbO_2	I intra-inter	0.0802	0.0618	0.0654	0.0951	68	10.7806	0.0000	0.0000
	I hom-nonhom	0.0898	0.0627	0.0747	0.1048	68	11.9046	0.0000	0.0000
	I rigth-left	0.0104	0.0957	-0.0126	0.0333	68	0.8991	0.3718	1.4870
	I short-long	0.1328	0.0523	0.1203	0.1454	68	21.1078	0.0000	0.0000
Variability HbO_2	I intra-inter	-0.0016	0.0352	-0.0100	0.0069	68	-0.3693	0.7131	2.8520
	I hom-nonhom	-0.0027	0.0338	-0.0108	0.0054	68	-0.6588	0.5122	2.0490
	I rigth-left	0.0004	0.0394	-0.0091	0.0098	68	0.0759	0.9397	3.7590
	I short-long	0.0027	0.0348	-0.0057	0.0110	68	0.6406	0.5239	2.0960

TABLE 3.5: Statistical analysis on the different indexes for Experiment 3 using HbO_2. T-tests against chance. Bonferroni correction was used.

Principal Components Analysis. As in the connectivity analysis for the familiarization phase of Experiment 1 we found 276 eigenvectors and its associated eigenvalues. From the eigenvectors we obtained the eigen-networks, while the eigenvalues describe the dynamic of the activation. The interesting aspect of the

		SumSq	DF	MeanSq	F	P
Strength: $I_{intra-inter}$ HbO_2	Condition	0.009	2	0.004	2.70821	0.07780
	Error (Condition)	0.071	44	0.002		
Strength: $I_{hom-nonhom}$ HbO_2	Condition	0.001	2	0.001	0.20313	0.81690
	Error (Condition)	0.117	44	0.003		
Strength: $I_{right-left}$ HbO_2	Condition	0.004	2	0.002	0.38946	0.67970
	Error (Condition)	0.204	44	0.005		
Strength: $I_{short-long}$ HbO_2	Condition	0.004	2	0.002	0.87263	0.42500
	Error (Condition)	0.100	44	0.002		
Variability: $I_{intra-inter}$ HbO_2	Condition	0.001	2	0.000	0.23261	0.79340
	Error (Condition)	0.051	44	0.001		
Variability: $I_{hom-nonhom}$ HbO_2	Condition	0.004	2	0.002	1.95721	0.15340
	Error (Condition)	0.040	44	0.001		
Variability: $I_{right-left}$ HbO_2	Condition	0.001	2	0.001	0.42899	0.65390
	Error (Condition)	0.059	44	0.001		
Variability: $I_{short-long}$ HbO_2	Condition	0.002	2	0.001	0.84258	0.43740
	Error (Condition)	0.044	44	0.001		

TABLE 3.6: Statistical analysis on the different indexes for Experiment 3 using HbO_2. 1-way ANOVA with condition as within subject factor.

applied analysis is that because we ran the PCA on the whole data set the modes of activation —or eigen-networks— are the same for the three conditions. Difference in the functional dynamic between them will be reflected on the eigenvalues.

As for the analysis of Experiment 1 the first eigenvectors capture the relevant variations of the data set, whereas the later ones are probably attributable to random noise (see Figures 3.15 and B.9). For HbO_2 we estimated with a significant level of 5% that the first 27 eigenvectors explained an amount of variance bigger that the expected when no temporal structure is present in the data, and the first 21 eigenvectors for Hb. The descriptive graph measures associated to these eigen-networks were also above the values obtained for the phase randomization simulations.

As for Experiment 1 we restricted our analysis to the first 4 eigenvectors. For HbO_2 the first four eigenvectors explained 33.0% of the total variation of the data (22.5%, 3.9%, 3.5% and 3.0% respectively). For Hb the first four eigenvectors explained 42.1% of the variation (33.8%, 3.2%, 2.7% and 2.4% respectively). The first four eigenvectors in the simulations explained on average 3.97% of the variance for HbO_2 (1.03%, 1.00%, 0.98% and 0.96% respectively), and 4.01% for Hb (1.04%, 1.01%, 0.99% and 0.97% respectively).

Functional networks topology: The first four eigen-netwoks or activation modes obtained for this experiment are remarkably similar to the ones of the previous experiment (see 3.16 and B.10). The first eigen-network has only positive weights

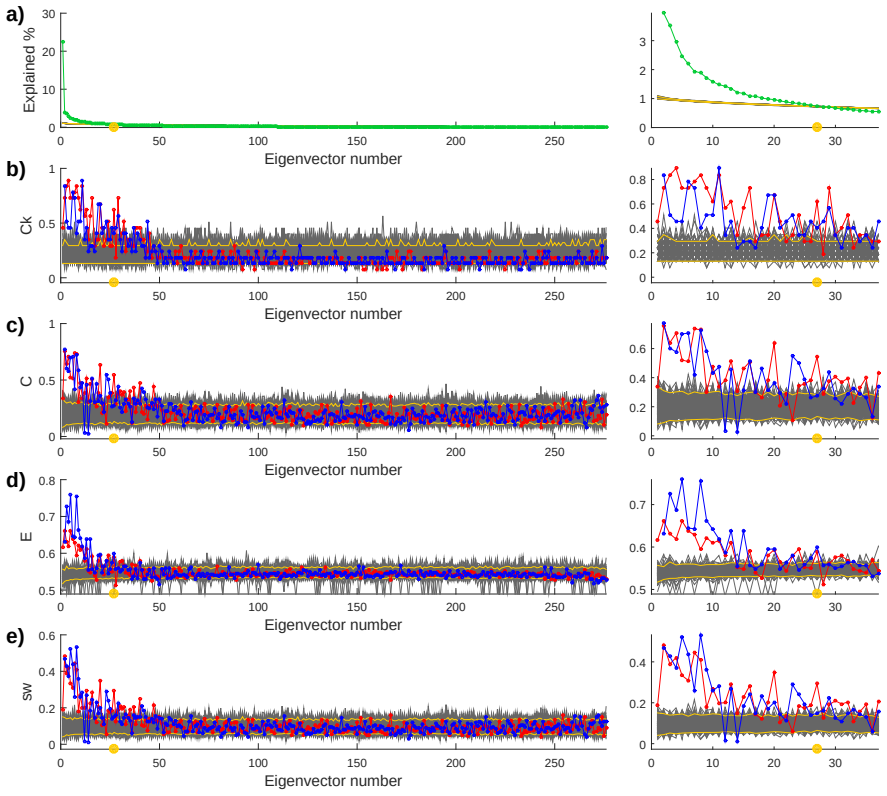


FIGURE 3.15: Phase randomization simulations for Experiment 3 using HbO_2. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

and lacks an efficient topological organization in terms of the chosen graph measures. The other eigen-networks have positive and negative weights. Eigen-network 2, as before, represents an activation mode in which more parietal areas increase their connectivity in opposition to more temporal areas. In eigen-network 3, connections involving frontal areas activate in opposition to connections involving more posterior regions. Finally, in eigen-network 4, the left and right hemisphere appear dissociated.

Dynamics of the activation: Difference across conditions in this experiment should appear in the eigenvalues. The time course of the eigenvalues for one subject in the three conditions are shown in Figure 3.11 with an illustrative purpose.

The eigenvalues associated to eigen-network 1 were strongly correlated with the mean correlation coefficient (R > 0.97, P < 1e-58), confirming it represents

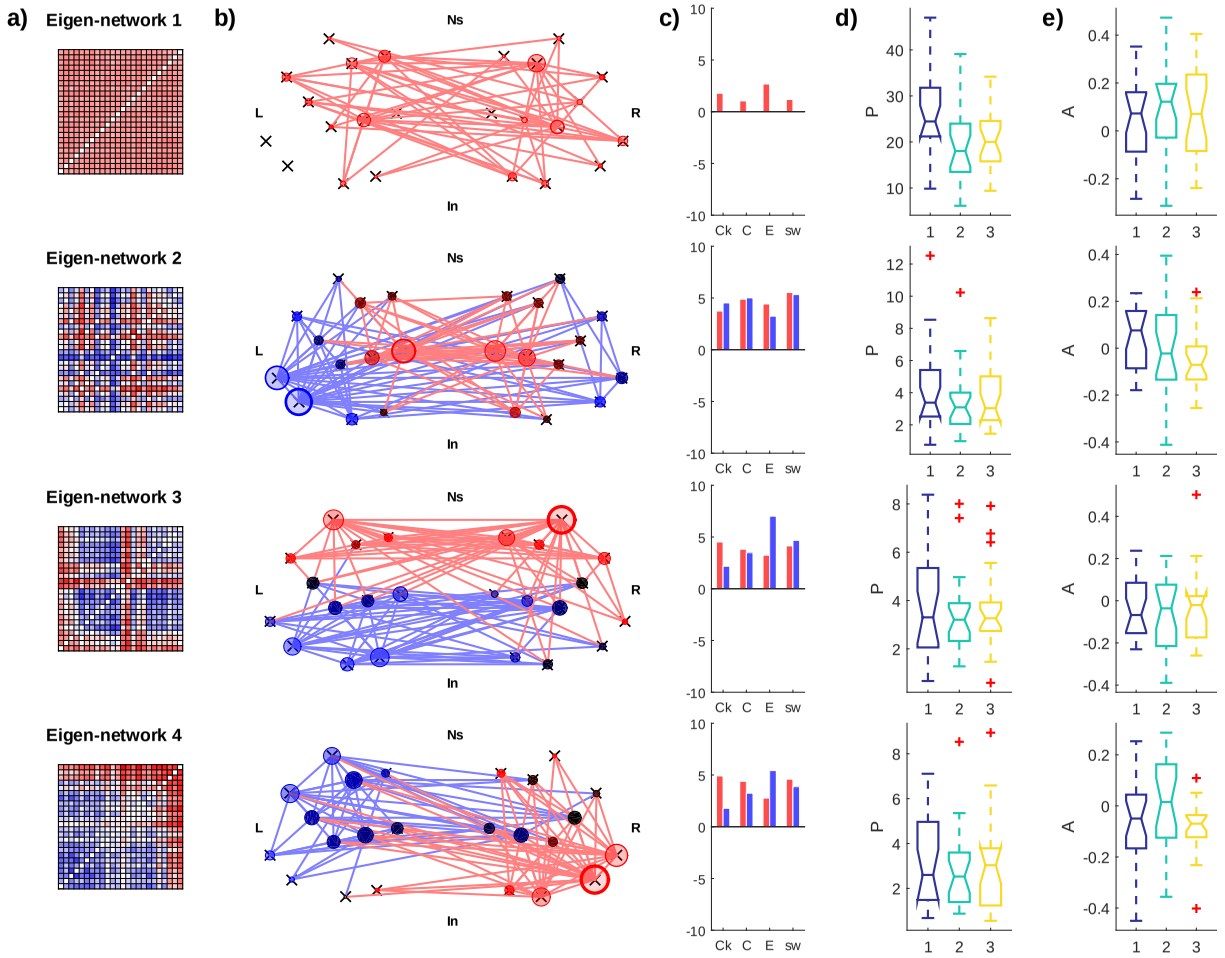


FIGURE 3.16: PCA analysis for Experiment 3 using HbO_2. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z-scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Explained variability (P) for each eigen-network per condition. TPs (blue), Rnd (green) and Sil (yellow). **e)** Asymmetry (A) per condition.

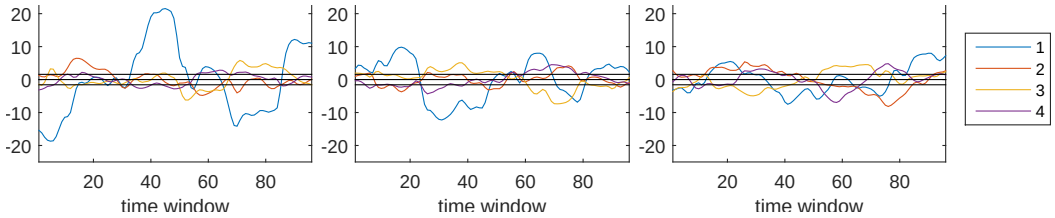


FIGURE 3.17: Eigenvalues along time for one subject for the first four components for Experiment 1. Results for HbO_2 are shown. From left to right TPs, Rnd and Sil conditions.

global connectivity changes.

In order to see how the different conditions modulate the functional connectivity dynamics we ran a 2-ways ANOVA using the variance explained by each component and the asymmetry in the distribution of the eigenvalues associated to each component as dependent variables.

Using HbO_2 to estimate functional connectivity, for the explained variance (P), we found a main effect of eigen-network (F(3,66) = 252.75, P < 0.0000), a main effect of condition (F(2,44) = 8.5045, P < 0.001), and a significant interaction (F(6,132) = 4.1199, P < 0.001) (see Table 3.7). Planed post-hoc Tukey-Kramer comparisons between components showed that the first component differed from all the others (P < 10e-8), but components 2-4 did not differ between them (P > 0.05) (see Table 3.9). When the different conditions were compared we found a significant difference between the TPs and Rnd condition (P < 0.0005), and a marginally significant difference between the TPs and Sil conditions (P = 0.047) (see Table 3.8). Planed post-hoc Tukey-Kramer comparisons also showed that the main effect of condition was driven by the the first eigen-network (for eigen-network 1, TPs vs. Rnd P < 0.005, TPs vs. silence P= 0.076; for all the other comparisons P > 0.05) (see Table 3.10).

		SumSq	DF	MeanSq	F	P
P	Condition	249.4	2	124.7	8.5	0.00080
HbO_2	Error (Condition)	645.1	44	14.7		
	Eigen-network	18369.2	3	6123.1	252.8	0.00000
	Error (Eigen-network)	1598.9	66	24.2		
	Condition: Eigen-network	467.4	6	77.9	4.1	0.00080
	Error (Condition: Eigen-network)	2496.0	132	18.9		
A	Condition	0.050	2	0.025	0.99327	0.37850
HbO_2	Error (Condition)	1.104	44	0.025		
	Eigen-network	0.604	3	0.201	7.513	0.00020
	Error (Eigen-network)	1.770	66	0.027		
	Condition: Eigen-network	0.176	6	0.029	1.10409	0.36340
	Error (Condition: Eigen-network)	3.516	132	0.027		

TABLE 3.7: Statistical analysis on the explained variance and asymmetry for Experiment 3 using HbO_2. 2-way ANOVA with condition and eigen-network as within subject factors.

When we compared the asymmetry (A) of the distribution of the eigenvalues we only found a main effect of eigen-network (F(3,66) = 7.5133, P < 0.001) (see

		Difference	StdErr	P	CI lower	CI upper
P	TPs-Sil	1.586	0.624	0.04712	0.018	3.150
HbO_2	TPs-Rnd	2.269	0.499	0.00045	1.015	3.520
	Rnd-Sil	-0.683	0.563	0.45834	-2.097	0.730

TABLE 3.8: Statistical analysis on the explained variance for Experiment 3 using HbO_2. Post-hoc multiple comparisons analysis between conditions. Turkey-Kramer correction was used.

Table 3.7). Pairwise Turkey-Kramer corrected comparison between the different eigen-networks uncovered that the difference was due to a higher asymmetry for the first eigen-network than for the third (P < 0.01) and for the fourth (P < 0.01) (see Table 3.9). Even if we did not find an effect of condition, we compared the asymmetry across conditions independently for each eigen-network. Applying Turkey-Kramer correction we found a significant difference between the TPs and the Sil condition for eigen-network 2 (P < 0.05) (see Table 3.10).

		Difference	StdErr	P	CI lower	CI upper
P	1-2	18.389	1.150	0.00000	15.196	21.580
HbO_2	1-3	18.752	1.065	0.00000	15.795	21.710
	1-4	19.332	1.186	0.00000	16.039	22.620
	2-3	0.363	0.276	0.56362	-0.404	1.130
	2-4	0.943	0.404	0.12101	-0.180	2.070
	3-4	0.580	0.333	0.32745	-0.345	1.500
A	1-2	0.075	0.029	0.07180	-0.005	0.160
HbO_2	1-3	0.110	0.030	0.00606	0.028	0.190
	1-4	0.118	0.027	0.00123	0.044	0.190
	2-3	0.035	0.029	0.63368	-0.047	0.120
	2-4	0.043	0.024	0.31008	-0.024	0.110
	3-4	0.008	0.028	0.99093	-0.068	0.080

TABLE 3.9: Statistical analysis on the explained variance for Experiment 3 using HbO_2. Post-hoc multiple comparisons analysis by eigen-network. Turkey-Kramer correction was used.

Using Hb we found similar results than with HbO_2. The 2-ways ANOVA on the percentage of variance explained revealed a main effect of eigen-network (F(3,66) = 529.11, P < 0.0000), of condition (F(2,44) = 4.2603, P < 0.05), and a significant interaction (F(6,132) = 3.6136, P < 0.005) (see Table B.6). The first eigen-network differed from all the rest (P < 1e-8) (see Table B.8). Between conditions the TPs condition was higher than the Sil condition (P < 0.05) (see Table B.7). Also in this case, this effect was driven by the first eigen-network, for which the TPs and Sil conditions were different (P < 0.05), whereas no significant differences were found for the rest of the eigen-networks (P < 0.05) (see Table B.9).

On the asymmetry we found a main effect of eigen-network (F(3,66) = 4.6077, P < 0.01), a marginally effect of condition (F(2,44) = 2.3471, P = 0.11), and not interaction (P > 0.05) (see Table B.6). The effect of eigen-network was due to a difference between eigen-network 2 with 3 (P < 0.05) and 2 with 4 (P < 0.05) (see Table B.7). Likewise for HbO_2 when we did pairwise Turkey-Kramer corrected comparisons between conditions independently for each eigen-network, and we

		Difference	StdErr	P	CI lower	CI upper
P	1: TPs-Sil	5.761	2.502	0.07646	-0.525	12.050
HbO_2	1: TPs-Rnd	7.471	2.107	0.00496	2.178	12.760
	1: Rnd-Sil	-1.710	2.149	0.70949	-7.107	3.690
	2: TPs-Sil	0.297	0.712	0.90870	-1.490	2.080
	2: TPs-Rnd	0.834	0.714	0.48373	-0.959	2.630
	2: Rnd-Sil	-0.537	0.682	0.71485	-2.251	1.180
	3: TPs-Sil	0.051	0.526	0.99484	-1.270	1.370
	3: TPs-Rnd	0.303	0.653	0.88869	-1.337	1.940
	3: Rnd-Sil	-0.252	0.570	0.89849	-1.685	1.180
	4: TPs-Sil	0.236	0.504	0.88692	-1.031	1.500
	4: TPs-Rnd	0.468	0.458	0.57004	-0.681	1.620
	4: Rnd-Sil	-0.232	0.576	0.91446	-1.679	1.210
A	1: TPs-Sil	-0.035	0.045	0.72967	-0.149	0.080
HbO_2	1: TPs-Rnd	-0.055	0.050	0.52570	-0.180	0.070
	1: Rnd-Sil	0.020	0.062	0.94244	-0.134	0.170
	2: TPs-Sil	0.088	0.033	0.03559	0.005	0.170
	2: TPs-Rnd	0.040	0.049	0.69982	-0.083	0.160
	2: Rnd-Sil	0.049	0.047	0.56645	-0.070	0.170
	3: TPs-Sil	0.003	0.046	0.99808	-0.113	0.120
	3: TPs-Rnd	0.028	0.049	0.83733	-0.094	0.150
	3: Rnd-Sil	-0.025	0.053	0.88416	-0.157	0.110
	4: TPs-Sil	0.025	0.042	0.82046	-0.081	0.130
	4: TPs-Rnd	-0.061	0.047	0.41902	-0.180	0.060
	4: Rnd-Sil	0.086	0.045	0.16400	-0.028	0.200

TABLE 3.10: Statistical analysis on the explained variance for Experiment 3 using HbO_2. Post-hoc multiple comparisons analysis by condition. Turkey-Kramer correction was used.

found a difference for eigen-network 2 between the TPs and Rnd conditions ($P < 0.05$) (see Table B.9).

Static functional connectivity. The Fisher transformed correlation matrix for the entire TPs, Rnd and Sil conditions and the associated functional networks using the top 20% correlations are presented in Figure 3.18. The correlation matrices show that correlations were globally stronger in the TPs condition. The networks for the three conditions are very similar and the channels involved are mainly over frontal areas. These networks do not present small-world properties. As for the previous experiment they resemble eigen-network 1 of the PCA analysis.

3.3.7 Discussions

The current experiment reproduced many of the findings of Experiment 1 for both analysis. The first analysis showed that intra-hemispheric connections were stronger than inter-hemispheric, that the functional connectivity between hemispheres was stronger for homologous regions, and that short range connections were stronger than long range connections. These features of the functional connectivity seem to be task independent and probably reflecting anatomical constraints. The PCA analysis provided eigen-networks remarkable similar to the ones individuated for Experiment 1, suggesting that their architecture is task independent as well.

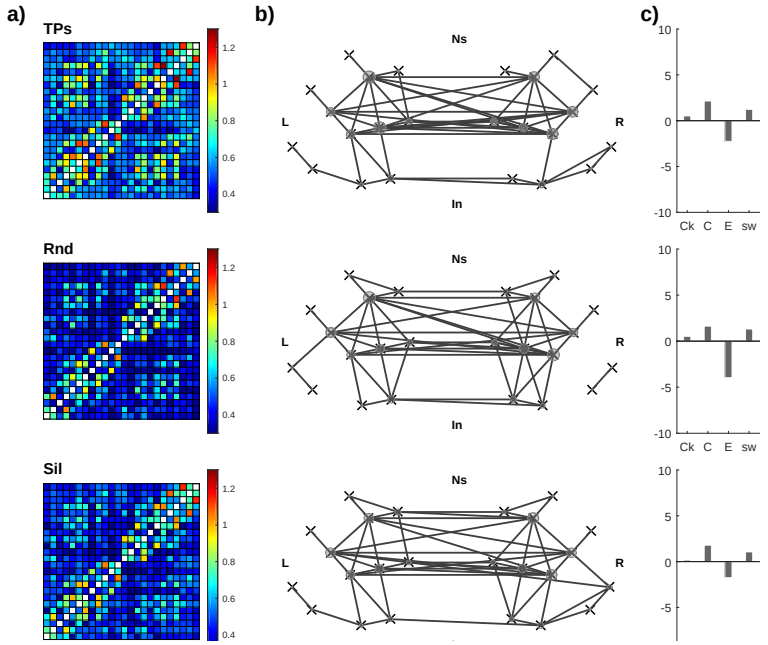


FIGURE 3.18: Static functional connectivity analysis for Experiment 3 using HbO_2. **a)** Fisher transformed correlation coefficient matrices. **b)** Networks obtained by thresholding (top 20 %) and binarizing the correlation matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z-scored graph measures for the network. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw).

The results are consistent with previous adults studies showing that functional networks are similar across tasks and mainly reflect the anatomical architecture of the brain (Cole et al., 2014). The task induces changes in functional connectivity, but the variations are much subtle (Cole et al., 2014; Davison et al., 2015; Hermundstad et al., 2013). Of course this does not mean that the functional connectivity we observed is a forthright image of structural connectivity. First, it is feasible that between two areas direct connections do not actually exist, but yet are functionally related. One explanation is that both are specialized in similar functions and their activity is coordinate by an indirect path —or simply by the stimulus, but in this case we would not expected a coordinate activity during rest. Second, it has to be considered that our connectivity measure is based on the BOLD signal. BOLD is an indirect measure of neural activity, thus the results may also reflect the organization and dynamic of the vascular system. Too little is now regarding brain vascularization in the neonatal brain to exclude this possibility.

Beyond the similarities across experiments and conditions we did find an effect of condition. The main effect of the task appeared in the modulation of the global strength and variation of the connections. The global connectivity was

consistently higher in the TPs than in the Rnd and Sil conditions, and the PCA analysis revealed that during TPs a bigger amount of variance could be explained by an eigenvector representing global connectivity changes. Both analysis evidenced that during the TPs condition the fluctuations in the global connectivity of the brain were stronger, with transient periods of a highly synchronized activity. A similar effect mainly involving frontal regions has been observed in adults and has been associated with attention and task performance (Braun et al., 2015; Bassett and Gazzaniga, 2011; Cole et al., 2013). The effect we observed was also stronger over frontal regions —note that our recording were limited mainly to fronto-temporal areas— which makes feasible that it is in fact related to the detection of the structure and extraction of the words, hence the results suggest a similar functional dynamic for the neonate brain and the adult brain.

Even if the stronger effect of condition was on the modulation of the global functional connectivity, it was not the only one. We also identified other aspects of the functional connectivity that were task dependent. The first analysis revealed some degree of modulation by condition for the indexes $I_{intra-inter}$ and $I_{short-long}$ on the strength (see Figure 3.14 and B.8); and in the PCA analysis the distribution of the eigenvalues for eigen-network 2 —which involves many inter-hemispheric connections— showed an effect of condition as well. Note that all these measures predicted task performance in Experiment 1.

The results of the current experiment do not allow us to completely disentangler if the correlations between functional connectivity and task performance arise from individual differences or also from changes induced by the task, but they give some inklings. The functional connectivity measures that correlate with task performance in Experiment 1 due to changes induced by the task, should be modulated by condition in Experiment 3. On the contrary, the functional connectivity measures correlated with task performance as pure consequence of individual differences, should not differ across conditions. $I_{intra-inter}$ and $I_{short-long}$ and eigen-network 2 showed both effects: predicted task performance in Experiment 1 and were modulated by condition in Experiment 3; suggesting that the task does induce a stronger synchronization between hemispheres and distant areas. On the contrary, stronger connectivity in the left hemisphere predicted performance in Experiment 1, but we did not observe a modulation of $I_{right-left}$ by condition in Experiment 3. This may indicate that the correlation with performance in Experiment 1 was consequence of the individual level of maturation of the left intra-hemispheric connections relative to the right connections. The modulations of the functional connectivity by condition, in line with results from Experiment 1 and consistent with the previous literature on brain development, supports a neural origin of the dynamic functional connectivity we observed.

Perhaps one of the main questions that remains open is the signification of the increase in the strength of global connectivity while neonates listen to the structured language. We hypothesize it is related with infants cognitive state, but

further research is needed. We have to denote that we cannot exclude the possibility that it reflects a physiological change induced by the task –to exclude this possibility heartbeat and respiration should be monitor. Nevertheless, considering that a similar effect has been reported in adults (Braun et al., 2015; Bassett and Gazzaniga, 2011; Cole et al., 2013) this possibility seems unlikely.

To summarize, the experiment revealed that the main architecture of the functional connectivity dynamic seems not to be affected by the task. However, a structured stream of syllables induced periods of high global functional connectivity, and a stronger synchronization between hemispheres and between distant areas within hemispheres, compared with silence and random syllables periods.

3.4 Functional connectivity while listening speech with prosodic cues. Experiment 2

Results from Experiments 1 and 3 showed that the main architecture of the functional networks seem to be task independent. In order to test it on another task and group of subjects we applied the exact same analysis than for Experiment 1 to Experiment 2. In Experiment 2 neonates had to use prosodic contours to extract the words, thus applying the same analysis to this experiment enables to evaluate which aspect of the functional connectivity dynamic that correlated with task performance in Experiment 1 were related with extracting the words, and which ones with the specific mechanism used for segmentation.

Based on the previous experiments we do not predict important differences in the eigen-networks of the PCA analysis. On the contrary, task performance should be predicted by the dynamic of the functional connectivity. But, because the properties of the stimuli that have to be tracked for segmentation are very different than in Experiment 1, differences could appear. While extracting the distribution of the syllables involves integration of information at a short temporal scale, segmenting based on prosodic contours implies processing supra-segmental information. A preponderant theory for speech processing states a dominance of the left hemisphere in processing speech over short time scales and of the right hemisphere for integrating information over longer time scale (Poeppel, 2003). Moreover, a superiority of the right hemisphere has been observed for processing supra-segmental information in young infants (Homae et al., 2006). Based on these considerations we expect task performance to correlate with stronger long range connections as in Experiments 1, because regardless of the segmentation cues used, extracting the words means forming some type of representation that presumably requires at least an immature phonological loop. However, it is an open question if differences between hemispheres will affect

neonates capacity to extract the words. Will the task induce a stronger synchronization in the right hemisphere?

A last consideration is that results may be less consistent. The differential activation for Part-words and Words was weaker in the prosody than in the TPs experiment, hence our measure of task performance is noisier.

3.4.1 Participants

Idem than for Experiment 2. See 2.4.1.

3.4.2 Stimuli

Idem than for Experiment 2. See 2.4.2.

3.4.3 Procedure

Idem than for Experiment 2. See 2.4.3.

3.4.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

3.4.5 Data Analysis

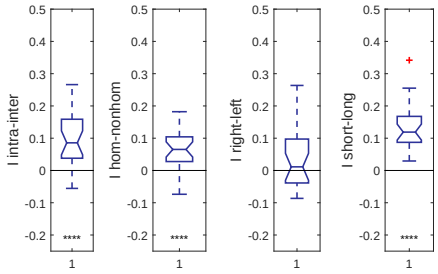
We applied the identical analysis than for Experiment 1 (see 3.2.5).

After the pre-processing, 26 out of the 40 infants of Experiment 2, could be included. The other 14 were excluded because of too many motion artifacts during the familiarization phase.

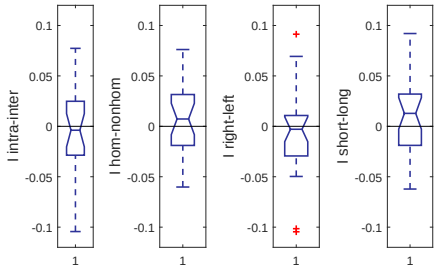
3.4.6 Results

The results for HbO_2 and Hb are presented below. Figures and tables of the Hb results are in the Appendix B.

Strength and variability of the connections. In Figures 3.19 and B.13 the indexes on the strength and temporal variability are presented for HbO_2 and Hb respectively. Two tails t-tests against chance level showed the same pattern of results than in the previous experiments (see Table 3.11 and Table B.10). Intra hemispheric connections are stronger than inter-hemispheric (for HbO_2, $P_{corr} <$ 0.0001; for Hb, $P_{corr} <$ 0.001); homologous connections are stronger than non-homologous (for HbO_2, $P_{corr} <$ 0.0001; for Hb, $P_{corr} <$ 0.0001), and short intra-hemispheric connections are stronger than long intra-hemispheric connections (for HbO_2, $P_{corr} <$ 0.0001; for Hb, $P_{corr} <$ 0.0001). Regarding the variability we only found a significant difference for $I_{short-long}$ in terms of Hb: short intra-hemispheric connections were less variable than long (P < 0.01).



(A) Strength



(B) Variability

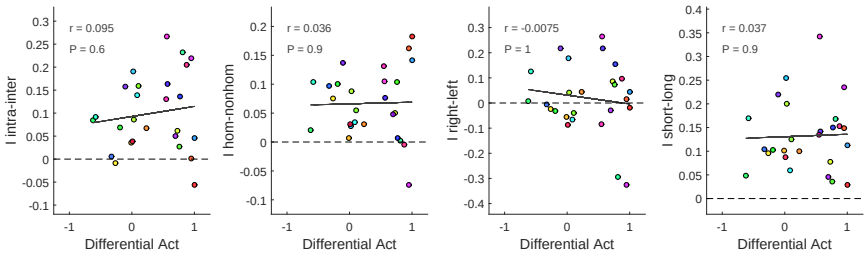
FIGURE 3.19: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 2. Connectivity was estimated using HbO_2. Asterisks bellow represent Bonferroni corrected P-values of the t-tests against zero.

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength HbO_2	I intra-inter	0.0999	0.0825	0.0666	0.1333	25	6.1787	0.0000	0.0000
	I hom-nonhom	0.0670	0.0595	0.0430	0.0910	25	5.7428	0.0000	0.0000
	I right-left	0.0196	0.1373	-0.0358	0.0751	25	0.7292	0.4727	1.8910
	I short-long	0.1324	0.0736	0.1027	0.1621	25	9.1747	0.0000	0.0000
Variability HbO_2	I intra-inter	0.0003	0.0509	-0.0203	0.0208	25	0.0272	0.9785	3.9140
	I hom-nonhom	0.0067	0.0338	-0.0069	0.0204	25	1.0141	0.3202	1.2810
	I right-left	-0.0100	0.0542	-0.0319	0.0119	25	-0.9429	0.3548	1.4190
	I short-long	0.0121	0.0372	-0.0029	0.0271	25	1.6631	0.1088	0.4350

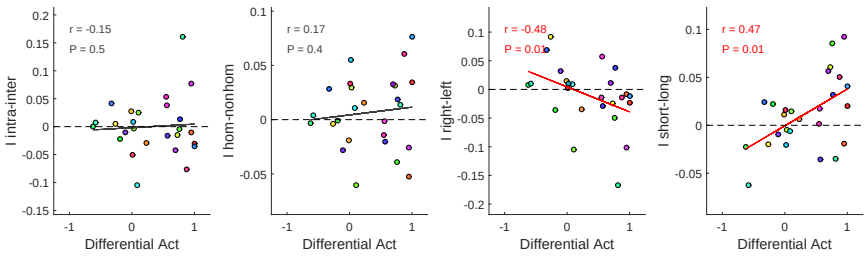
TABLE 3.11: Statistical analysis on the different indexes For Experiment 2 using HbO_2. T-tests against chance. Bonferroni correction was used.

When we correlated the asymmetry indexes with the differential activation for part-words and words during test trials, we did not find any significant correlation with the indexes on the average strength but we did find significant correlations with the right - left, and short - long indexes on the variability. $I_{right-left}$ was negatively correlated with the differential activation for HbO_2 (R = -0.4753, P < 0.05) and marginally fo Hb (R = -0.3655, P = 0.0663); and $I_{short-long}$ positively correlated for both HbO_2 (R = 0.4722, P < 0.05) and Hb (R = 0.5451, P < 0.01).

Principal Components Analysis. The phase randomization simulation showed that the first components explained more variability of what would have been expected when the temporal structure in the data is missing. We estimated, with a 5% confident level, that this applies to the first 17 eigenvectors for HbO_2 (see



(A) Strength



(B) Variability

FIGURE 3.20: Spearman correlations between task performance (differential activation during test blocks) and functional connectivity for Experiment 2 using HbO_2. The differential activation for part-words and words during test blocks (x-axis) is plotted against the different indexes estimated during the familiarization phase (y-axis): $I_{inter-intra}$, $I_{hom-nonhom}$, $I_{right-left}$, $I_{short-long}$. Each dot represent a subject.

.

Figure 3.21), and to the first 14 for Hb (see Figure B.15). Likewise in the previous experiments, for the first eigen-networks, the descriptive graph measures for the real data were also higher that the ones obtained from the phase randomized data. As in Experiment 1 we restricted our analysis to the first four eigenvectors. For HbO_2 they explained 32.4% of the variability (22.3%, 3.6%, 3.3% and 3.2% respectively), and for Hb 39.6% (30.7%, 3.6%, 3.0% and 2.4% respectively). In the phase randomization simulations the first four components explained 6.18% of the variance for HbO_2 (1.63%, 1.56%, 1.51% and 1.47% respectively), and 6.20% for Hb (1.63%, 1.56%, 1.52% and 1.48% respectively).

Functional networks topology: The first four eigen-networks are presented in Figure 3.22 for HbO_2, and in Figure B.16 for Hb. As in the previous experiments the first eigen-network has only positive weights, represents global changes in the functional connectivity, and explains most of the variance in data. The next eigen-networks present positive and negative weights and show an efficient topological organization in terms of all the chosen descriptive graph measure. The structure of the eigen-networks fairly resembles the structure found for the other experiments, in particular for eign-network 2. Eigen-networks 3 and 4 show activation modes in which anterior and posterior, and left and right areas are in opposition, but the division is slightly different than in previous experiments.

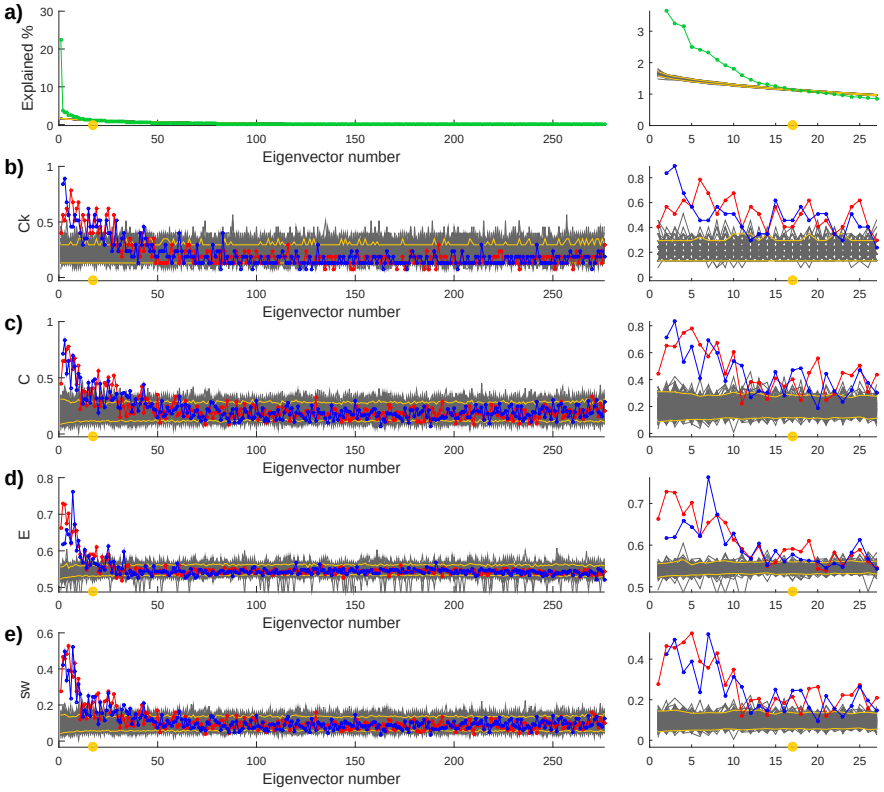


FIGURE 3.21: Phase randomization simulations for Experiment 2 using HbO_2. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

Dynamics of the activation: The eigenvalues associated to eigen-network 1 were strongly correlated with the mean correlation over all connections pairs along time (for HbO_2 R > 0.995, P < 1e-94 for all subjects; for Hb R > 0.998, P < 1e-118 for all subjects), confirming that it represents global changes in connectivity.

We did not find significant correlations between the dynamic of the activation during familiarization and the differential activation during test blocks. No effect or trend was observed neither for the explained variance by eigen-network 2 — which was positively correlated with task performance in Experiment 1; nor for the asymmetry of eigen-netwok 4—which was negatively correlated.

Static functional connectivity. The Fisher transformed correlation matrix for the entire familiarization and the functional network built using the top 20% correlations are presented in Figure 3.23. As in the previous experiments it involves

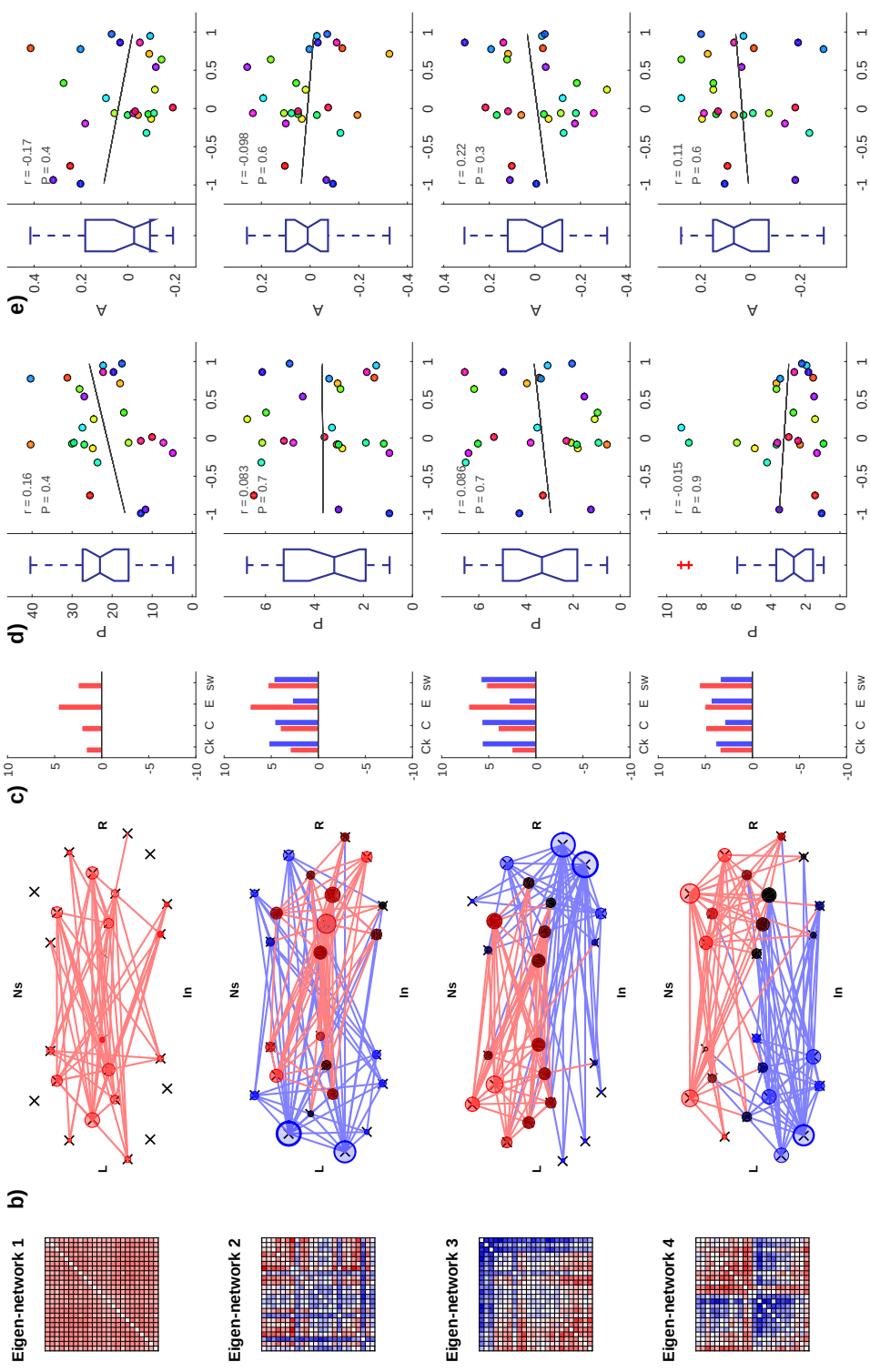


FIGURE 3.22: PCA analysis for Experiment 2 using HbO_2. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Spearman correlations between the explained variability (P) for each eigen-network and the differential activation for Part-words and Words. **e)** Spearman correlations between the asymmetry (A), for each eigen-network and the differential activation for Part-words and Words.

mainly frontal channels and reassembles eigen-network 1.

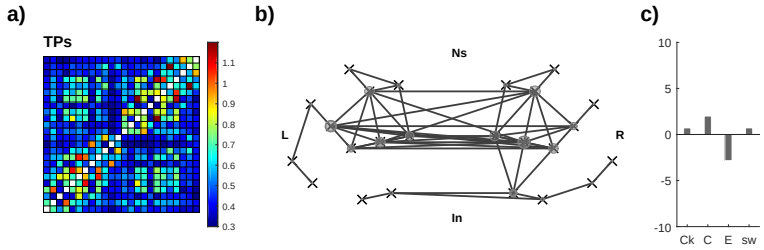


FIGURE 3.23: Static functional connectivity analysis for Experiment 2 using HbO_2. **a)** Fisher transformed correlation coefficient matrix. **b)** Network obtained by thresholding (top 20 %) and binarizing the correlation matrix. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the width of its line the centrality of the node. **c)** Z-scored graph measures for the network. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw).

3.4.7 Discussions

The main structure of the functional connectivity was similar to previous experiments as we expected. In the first analysis connections were stronger within hemispheres, between homologous regions, and in closer areas within an hemisphere. The eigen-networks obtain in the PCA analysis in general terms resemble the ones found for the previous experiments and as before presented small-world properties.

Regarding how task performance was predicted by the functional connectivity dynamic during the familiarization, results were less clear than for Experiment 1. For the first analysis, the measures correlated with task performance were based on differences in variability. The $I_{right-left}$ was negatively correlated with performance, which implies either more stable right connections or more variable left connections for good performers; whereas the $I_{short-long}$ was positively correlated with task performance, implying that infants performing better had either more stable long connections or more variable short connections. Note that task performance in Experiment 1 was predicted by the same indexes but in terms of the strength. The interpretation of the results is not straightforward, nevertheless, we can expose a possibility in line with the proposal of the left and the right hemispheres specialized in speech processing at different time scales (Poeppel, 2003). Better performance in Experiment 1 is achieved by infants with stronger connections in the left hemisphere and stronger long range connections. Extracting the words in Experiment 2 should rely more on the right hemisphere, which is in general more matured at birth (Leroy et al., 2011). This is maybe way difference between subjects in performance appeared in term of how stable are

the connections —better performers have more stable right and long range functional connections. Of course, this remains just a conjecture; further research on the functional development of the brain and the processing of distributional and prosodic cues is need in order to make valid conclusions. Moreover, as we said in the introduction, the differential activation for part-words and words in Experiment 2 was not so strong as in Experiment 1, making the correlations more susceptible noisier to noise in the data.

3.5 Chapter Discussions

In the current chapter we described the dynamic of functional connectivity in the neonatal brain during speech segmentation tasks and rest. To our knowledge this is the first time that this is studied in infants, and the first time this type of analysis is performed in fNIRS recordings. We used a sliding time window approach and performed two analysis. A simple comparison of the strength and variability of the connectivity between different connections types; and an analysis adapted from Leonardi et al., 2013, in which PCA is applied to the variations in connectivity. Both methods are suitable for the study of dynamic aspects of the functional connectivity in infants using fNIRS recording.

The essential aspect of the PCA method is that it does not focus on the stronger connections at each time window, but it extracts networks based on groups of connections that covary —modes of variation of the connectivity. Two aspects are crucial. First, it does not restrict to patterns obtained from the stronger connections, which is important because the average correlation between two areas is probable a consequence of numerous factors. A second order analysis as this one neglects the “baseline level" of correlation and focus on the variations relative to it. Second, the method makes possible to easily identify partially overlapping networks, which seems fundamental considering that a same area may be involved in multiple functions.

We identified properties of the functional connectivity that were stable and independent of the task. The activity in homologous and in spatially closed areas was more synchronized, a result that reproduces the finding of previous functional connectivity analysis that assumed stationarity (Homae et al., 2010; Fransson et al., 2011; Perani et al., 2011). With the PCA analysis we obtain reproducible modes of variation of the functional connectivity. The networks had a characteristic topology with small-world properties and consisted on groups of connections with opposite behaviours compatible with a feasible functional organization of the brain: left opposite to right, frontal opposite to posterior, homologous connections between parietal regions opposite to connections between more temporal areas. Remarkably, our measures over frontal areas do not appear disconnected and show a strong activity. All together the results suggest that since birth the brain presents a well stablish functional organization including also frontal areas.

Our result are also consistent with homologous areas being involved in similar functions.

The most novel aspect of our results regards the dynamic properties of the functional connectivity. We observed that as in adults, functional connectivity variations are dominated by periods of globally high and low connectivity. Interesting these variations were higher while neonates were listening to a structured stream, a result that is compatible with adults studies showing that global connectivity is affected by factors like attention (Braun et al., 2015; Bassett and Gazzaniga, 2011; Cole et al., 2013). It is promising the study of changes in global connectivity relative to the cognitive state during development, because it could reveal interesting features of infants cognition.

As I have extensibility discussed in previous sections the effect of the task seems not to be restricted to the changes in the global connectivity. Even if the general functional architecture remained stable we did observe differences across tasks. In brief, our results suggest two main things. First that segmenting the speech and extracting the words is associated with stronger long range within hemispheres connections, independently of the segmentation cue present in the stimuli. Second, the results are consistent with a hemispheric specialization for speech processing present from birth, with the left hemisphere dominant in processing features at a short time scale, and the right hemisphere in the processing of supra-segmental information as prosody.

We demonstrated that as in adults (e.g. Park et al., 2012; Bassett and Gazzaniga, 2011; Allen et al., 2014), functional connectivity is not static but changes along time. Understanding the patters of variation governing this phenomenon becomes central, and we believe its study in infants should be encouraged.

To finalize, we performed all the analyses independently on both HbO_2 and Hb. Correlations coefficient were in general stronger for Hb, but we obtained pretty consistent results between the two signals.

Chapter 4

Short term memory and serial order effects in neonates

4.1 Are serial position effect relevant for language acquisition and language structure?

The temporal nature of language requires encoding sequential information; in spoken language, syllables are sequentially organized into words, words into phrases, and phrases into sentences. For example, learning that the word banana refers to the concept of a banana requires that infants encode the sequence of syllables that comprise the word. They must learn that, not only does the word consist of three syllables /ba/, /na/, and /na/, but also that those three syllables are arranged in a specific order. At the sentence level, the sentence the dog bites the boy has an entirely different meaning from the sentence the boy bites the dog, while dog boy the bites the is completely ungrammatical and meaningless; even though each sentence is constructed from the same words, the position of those words alters the meaning. In the current chapter we investigate whether a signature constraint of sequential processing in adults— an enhanced encoding of sequence edges— is evident from birth and constrains language processing.

Since language relies on the ability to encode sequential information, sequential processing constraints are related to the way in which language is processed. Extensive research shows that a fundamental constraint in processing sequential information is that edges of a sequence are encoded more precisely than internal positions. This has been demonstrated in two distinct, but related ways. First, the items at the edges of a sequence are better recalled than items in the middle, a phenomenon known as the serial position effect (Ebbinghaus, 1885). This enhanced memory for items at the edges is robust, with the same pattern emerging across many domains (Gupta et al., 2005, for a review see Hurlstone, Hitch, and Baddeley, 2014). Second, when encoding the order of items in a sequence, only the edge positions appear to be encoded precisely, while all other positions appear to be encoded relative to the edges (Endress, Nespor, and Mehler, 2009; Henson, 1998). In a sequence, such as ABCDE, A is encoded precisely as the first position and E as the last position, while BCD are encoded less precisely, only relative to

the edge positions (e.g., B is one position after the first position or three positions before the last position). Edge positions are better encoded than internal positions (Henson, 1998).

Evidence that this enhanced encoding of edges influences language processing comes from both cross-linguistic research and artificial grammar learning experiments. Across languages, linguistic cues are typically edge based, with determiners, bound morphemes, and linguistic stress position appearing with respect to the edge of a word (Greenberg, 1957; Hayes, 1995). Artificial grammar learning experiments demonstrate an enhanced encoding of items located at the edges of sequences (Endress, Nespor, and Mehler, 2009). Adults extract a positional-based regularity (e.g., an immediate repetition of two syllables in a seven syllable sequence) when that regularity is at the edge of the sequence but not when it is internal (Endress and Mehler, 2009a). Adults fail to extract a structural regularity —an AxC pattern with words, in which A always predicts C with an irrelevant syllable separating them— from a continuous stream, but succeed when the words are separated by an imperceptible 25ms pause (Peña et al., 2002). When the boundaries were marked by another cue, even if it was very subtle, the regularity at the edge positions was generalized.

Combined, the existing research suggests that when processing speech, the reliance on sequential information privileges encoding of the edges. Little research, however, has examined how sequential information is encoded early in development and whether there are constraints on sequential processing in infancy that might influence early language acquisition. Some studies have demonstrated long-term memory serial position effects in 3- and 6-month-olds using small visual sequences (Gulya et al., 2001; Gulya et al., 1998). A recent study (Benavides-Varela and Mehler, 2015) demonstrated that 7-month-olds show a more precise encoding of the syllables at the edges of a word than of syllables in the middle of a word; they detected a change when the edge syllables of a 5-syllabic word switched positions but not when two of the internal syllables switched positions. While this suggests that an enhanced encoding of sequence edges emerges during the first year of life, it remains an open question whether this bias is an inherent constraint on sequential processing. At this age, infants are already sensitive to some edge-based linguistic regularities (Gervain et al., 2008a; Seidl and Johnson, 2006; Thiessen and Saffran, 2007) and exposure to such robust cues may tune infants to an enhanced encoding of sequence edges.

4.2 Neonates encode better the edge syllables of a sequence. Experiment 4

In Experiment 4 we tested whether the edge bias is a processing constraint that is present at birth, prior to extensive linguistic experience. Specifically, we ask if neonates encode positional information from six-syllabic sequences and if they

encode some positions (i.e., syllables at the edges) more precisely than others (i.e., syllables in the middle), in accordance with constraints on general sequential processing found in adults.

Newborns are capable of remembering bisyllabic sequences after brief familiarizations shortly after birth (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012), but no studies have examined whether newborns remember information from longer multisyllabic sequences. We familiarized newborn infants to a six-syllabic sequence and then switched either the edge or internal syllables to examine whether they detected the change. As in Benavides-Varela et al., 2011; Benavides-Varela et al., 2012 and our previous experiments (Experiments 1 and 2) we measure fNIRS habituation / novelty responses. We familiarized neonates to a six-syllabic sequence and examined if they could detect a change when either the two edge syllables switched position (Edge Switch condition) or two internal syllables switched position (Internal Switch condition). Switching the syllables, rather than replacing them with novel syllables allowed us to determine whether infants were able to encode, not just the syllabic identity, but the syllables' positional information within a sequence as well. If neonates encoded all syllables in the sequence equally, they should detect a change on both conditions, that will manifest in an increase in the hemodynamic response in both conditions. However, if neonates better encode the edge syllables they would detect the difference only in the Edge Switch condition.

4.2.1 Participants

All participants were full-term neonates born to Italian-speaking mothers, with Apgar score ≥ 7 in the first minute and ≥ 8 in the fifth minute, diameter of head ≥ 33.0 cm, and no cefalhematomas. The Edge Switch condition included 16 participants (6 females; mean age 2.75 days, range 2–4 days; mean gestational age 38.9 weeks, SD 1.1 week; mean weight 3.475 kg, SD 0.336 kg), and the Internal Switch condition included 16 participants (10 females; mean age 3.1 days, range 2–4 days; mean gestational age 38.9 weeks, SD 1.1 week, mean weight 3.382 kg, SD 0.359 kg), who provided data without motion artifacts from at least two of the six of the test blocks. Additional infants were tested but excluded from the final analyses because of motion artifacts during more than four of the test blocks (Edge Switch condition n = 11, Internal Switch condition n = 8), failure to complete the experiment (Edge Switch condition n = 3, Internal Switch condition n = 5), a poor signal due to thick hair (Edge Switch condition n = 6, Internal Switch condition n = 2), or computer error (Edge Switch condition n = 1). All newborns were recruited from the nursery at Hospital, Azienda Ospedaliera Santa Maria della Misericordia, in Udine, Italy. Parents provided informed consent. The Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

Edge Switch		Internal Switch	
Fam	**Test**	**Fam**	**Test**
simebutalɛfo	fomebutalɛsi	simebutalɛfo	simetabulɛfo
fomebutalɛsi	simetabulɛfo	simetabulɛfo	simebutalɛfo
nɛkalisorevu	vukalisorenɛ	nɛkalisorevu	nɛkasolirevu
vukalisorenɛ	nɛkalisorevu	nɛkasolirevu	nɛkalisorevu
gamɛzibekotu	tumɛzibekoga	gamɛzibekotu	gamɛbezikotu
tumɛzibekoga	gamɛzibekotu	gamɛbezikotu	gamɛzibekotu
nelokisubɛma	malokisubɛne	nelokisubɛma	nelosukibɛma
malokisubɛne	nelokisubɛma	nelosukibɛma	nelokisubɛma
ponivelagusɛ	sɛnivelagupo	ponivelagusɛ	ponilavegusɛ
sɛnivelagupo	ponivelagusɛ	ponilavegusɛ	ponivelagusɛ
boketasɛluma	maketasɛlubo	boketasɛluma	bokesɛtaluma
maketasɛlubo	boketasɛluma	bokesɛtaluma	boketasɛluma

TABLE 4.1: Stimuli for Experiment 4. The test word was created exchanging the first and last syllables for the Edge Switch condition and the intermediate syllables in the Internal Switch condition. Different pairs of familairzation and test words were used. Infants were randomly assigned to conditions and a pair of familiarization and test words were randomly assigned.

4.2.2 Stimuli

All sequences consisted of six different consonant-vowel (CV) syllables. The sequences were synthesized using the it4 Italian female voice of the MBROLA diphone database (Dutoit et al., 1996), with phoneme duration of 150 ms and a constant pitch of 200 Hz. Sequences were continuous with no pauses between the syllables. All test sequences were generated by switching the position of two syllables from the familiarization sequences. In the Edge Switch condition, the first and the last syllables switched position. In the Internal Switch condition the third and fourth syllables switched position. In order to avoid that results were driven by specific properties of the stimuli, twelve different pairs of familiarization and test sequences were generated per condition, and infants were randomly assigned the condition and the pair of sequences that were used during familiarization and test blocks (see Table 4.1). Note that each sequence used during the familiarization for one infant was used for the test phase for another (e.g. one infant in the Edge Switch condition heard simebutalɛfo during familiarization blocks and fomebutalɛsi during the test blocks while another infant heard fomebutalɛsi during familiarization blocks and simebutalɛfo during the test blocks).

4.2.3 Procedure

We implemented a between subjects design The experiment consisted of six sets of familiarization and test blocks (See Figure 4.1). During each familiarization

block the same sequence was repeated 20 times, separated by a silence of 0.5–1.5 s. During each test block the test sequence was repeated five times, separated by a silence of 0.5–1.5 s. The total duration was 56 s for each familiarization block, and 14 s for each test block. All blocks were separated by periods of silence of random lengths between 25 and 30 s to allow the hemodynamic response to return to baseline. The total duration of the experiment was 13 minutes.

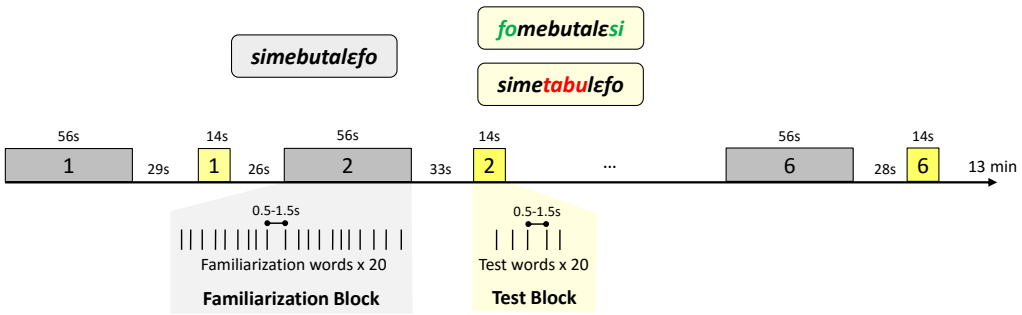


FIGURE 4.1: Schematic representation of the protocol for Experiment 4. The experiment consisted on a series of six familiarization and test blocks. During familiarization blocks a word was presented 20 times. During test blocks half of the infants heard a word created by exchanging the first and last syllables, and the other half a word created exchanging the two intermediate syllables.

4.2.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

4.2.5 Data Analysis

The analysis was essentially the same described for Experiments 1 and 2, with some differences in the pre-processing. Data of the current experiment were analysed before data from Experiments 1 and 2, thus modifications in the pre-processing of the latest experiments were introduced with the aim of improving the quality of the data. The main steps of the analysis —pre-processing, calculation of the HRF, data rejection, and statistical analysis— and the differences respect to the analysis used in Experiments 1 and 2 are described below.

Pre-processing. Data were analysed using custom functions in MATLAB 2012b (MATLAB and Statistics Toolbox Release 2012b, n.d.). The pre-processing involved the following steps:

1. *Identification of periods of saturation or low signal to noise ratio.* We marked saturated samples as those with a light absorption of less than 1% of the total light; and samples with low signal to noise ratio as those with a ratio

between the standard deviation and the mean greater than 5 in a moving window of 5 s.

2. *Detection of motion artifacts by fast changes in the signal.* We used the *hmrMotionArtfactsByChannel* function of the Homer2 NIRS package. The function identifies samples in a moving time window of length tMotion ± tMask as motion artifacts if there is a change greater than a threshold, amp_thresh, or if any value in the time window is bigger than std_thresh standard deviations. Unlikely Experiments 1 and 2 we did no used the modified version of the function, thus the threshold for the fast change was not determined for each channel and subject, but it was fixed. We used the z-scored data per subject in order to make the parameters comparable with other experiments, equipment, etc. The parameters we used were: tMotion = 0.5s, tMask = ± 0.3s, amp_thresh = 0.5 and std_thresh = 4.

3. *Correction of motion artifacts by spline interpolation.* We corrected motion artifacts using the spline interpolation algorithm, by applying the hmrMotionCorrectSpline function with p_spline = 0.99 (Scholkmann et al., 2010). This algorithm independently fits each artifact using cubic spline interpolation and subtracts the fit of the motion artifact from the signal.

4. *Conversion of the intensity to optical density.* We converted the intensity to optical density using the *hmrIntensity2OD* function of the Homer2 NIRS package, as we did for Experiments 1 and 2 (see 2.3.5).

5. *Calculation of the relative changes in HbO_2 and Hb.* To do so we used the *hmrOD2Conc* function of the Homer2 NIRS package in identical way than for Experiments 1 and2 (see 2.3.5). Afterwards we linear detrended the entire time series.

6. *Re-detection of motion artifacts.* As in the pre-processing used for Experiments 1 and 2 some motion artifacts —usually too long or too strong— are not properly corrected and data has to be rejected. We applied the motion artifact detection algorithm again, but on the HbO_2 with the following parameters: tMotion = 0.5, std_thresh = 4, amp_thresh = 0.08. We considered there was a motion artifact if fast changes were detected in at least 12 out of the 24 channels for HbO_2. Notice that we did not check for motion artifact again in Hb, because initially we did not analyse it.

HRFs calculation. We calculate the HRFs in the same way than for Experiments 1 and 2 with small variations in the parameters. We band-pass filter the pre-processed data between 0.02 Hz and 0.80 Hz. To extracted the HRFs from the HbO_2 and Hb time series for each Test Block we cut from -5s to +29s from

the onset of the stimuli. We used the mean value in the period [-5s-0s] and [24s-29s] to calculate a linear baseline trend that was removed from the signal. Finally, and after data rejection, we calculated an average HRF per infant per channel per condition, that we used for statistical analysis.

Data rejection. An HRF for a block and channel was excluded if during that period we detected either (1) saturation of the signal, (2) low signal to noise ratio, or (3) not properly corrected motion artifacts. Moreover we visually inspected the data that appeared as outliers but that had not been identified as motion artifact by the algorithm, and we compared to annotations made during the experiment and the video recording. If motion artifacts were identified, data was manually rejected. Considering all the infants tested (including infants who were ultimately excluded for not providing enough good data), in the Edge Switch condition, 49% out of the total HRFs, were automatically rejected and 8% manually rejected; whereas in the Internal Switch condition, 50% of the HRFs were automatically rejected and 6% manually rejected. Infants were included in the analysis only if they contributed at least with two good test blocks per condition. Each of the included infants contributed an average of 3.81 blocks (SD = 0.91) in the Edge Switch condition, and 3.50 (SD = 1.15) in the Internal Switch condition.

Statistical Analysis. As for Experiments 1 and2 we performed a cluster based permutation analysis and a more traditional mean activation analysis.

Cluster Based Permutation Analysis. We used the Cluster Based Permutation Analysis described in section 2.3.5 to compare the HRFs for test blocks during the Edge and Internal Switch conditions. We run the analysis in the HRFs obtained from the pre-processing during the period [-5s, +29s], hence lasting 34 s. Before running the analysis we smooth the data by down sampling to 1Hz, therefore we obtained 816 pairs of data points to compare (24 channels x 34 samples). We used a two tails two sample t-tests and 0.05 as threshold p-value to select the pairs of samples that constitute the clusters candidates. We considered two pairs of samples temporally adjacent if they were consecutive (time difference of 1s), and spatially adjacent if they were at a distance < 3cm (See Table 2.2). We run 1000 randomizations in the permutation analysis to obtain the Monte Carlo p-value.

Mean Activation Analysis. We did the Mean Activation Analysis to validate the results form the Cluster Based Permutation Analysis. Furthermore, because we consider the same regions of interest than in Experiments 1 and 2, we can compare activation patterns across experiments. We calculate the mean change in HbO_2 and Hb during the test blocks in the time window [0s, +24s], and we used it as dependent measure in a 3-ways ANOVA. We included hemisphere (left/right), and region (frontal/ temporal/ parietal) as within subject factors, and condition

(Edge/ Internal) as between subjects factor.

4.2.6 Results

Results from both analyses are presented below. Figures and Tables for results based on Hb are presented in the Appendix C.

Cluster Based Permutation Analysis. The analysis revealed a greater hemodynamic response for the Edge Switch condition than for the Internal Switch condition in both hemispheres and mostly in tempo- frontal areas. For HbO_2 we found a cluster in the left hemisphere including channels 3, 5, 6, 7, 9 and 11, within the time window [6s - 16s] from the onset ($P_{HbO_2}cluster_1 < 0.01$); and one cluster in the right hemisphere including channels 13, 14, 15, 17, 18, 19, 21 and 22, within the time window [4s - 13s] ($P_{HbO_2}cluster_2 < 0.01$) (see Figure 4.2). For Hb we found a significant cluster in the right hemisphere including channels 13, 15, 16, 18 and 21 within the time window [19s - 26s] ($P_{Hb}cluster_1 < 0.01$) (see Figure C.1), and two marginally significant clusters in the left hemisphere, one including channels 3, 5, 6 and 7 in the time window [12s - 16s] ($P_{Hb}cluster_2 = 0.054$), and the other comprising the channels 9 and 11 within the time window [23s - 26s] ($P_{Hb}cluster_3 = 0.079$).

Mean Activation Analysis. The analysis confirmed the previous results. For HbO_2 we found a main effect of condition (Edge / Pause) (F(1,15) = 19.646, P < 0.001), a marginally significant effect of region (frontal/ temporal/ parietal) (F(2,30) = 3.0744, P = 0.0610), but not effect of hemisphere, and not significant interactions (P > 0.05) (see Figure 4.3 and Table 4.2). A post-hoc multiple comparison analysis, Turkey-Kramer corrected, revealed that the effect of region was due to a stronger activity in frontal than in temporal areas (P = 0.0679). For Hb we found a main effect of condition (F(1,15) = 19.521, P < 0.001), and not other significant effects or interactions (see Figure C.2 and Table C.1).

HbO_2	SumSq	DF	MeanSq	F	P
Condition	0.04066	1	0.04066	19.64602	0.00050
Error (Condition)	0.03105	15	0.00207		
Hemisphere	0.00093	1	0.00093	1.44515	0.24790
Condition: Hemisphere	0.00043	1	0.00043	0.67264	0.42500
Error (Hemisphere)	0.00964	15	0.00064		
Region	0.00303	2	0.00152	3.07437	0.06100
Condition: Region	0.00035	2	0.00017	0.35257	0.70580
Error (Region)	0.01479	30	0.00049		
Hemisphere: Region	0.00062	2	0.00031	0.63160	0.53870
Condition: Hemisphere: Region	0.00037	2	0.00018	0.37632	0.68960
Error (Hemisphere: Region)	0.01464	30	0.00049		

TABLE 4.2: Statistical analysis for Experiment 4. 3-ways ANOVA on the mean $[HbO_2]$ as dependent variables. Hemisphere (left/ right) and region (frontal/ temporal/ parietal) are within subject factors, and condition (Edge/ Internal) is a between subject factor

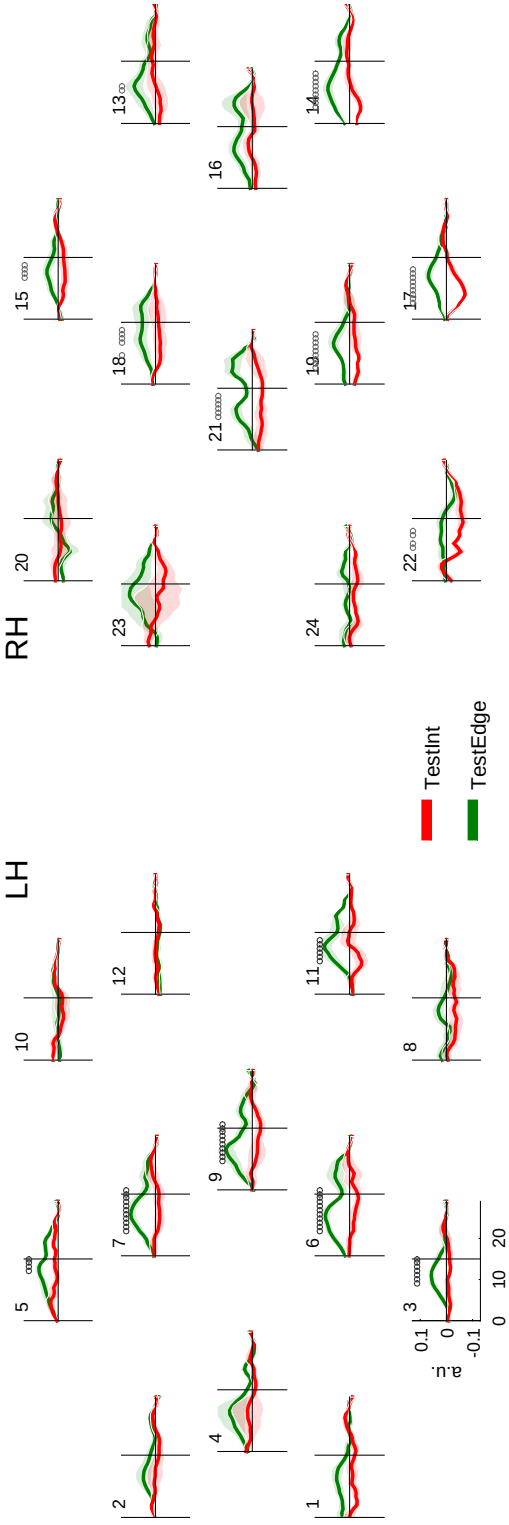


FIGURE 4.2: Cluster based permutation analysis for Experiment 4 using HbO_2. HRFs for the Edge Switch condition (green) and Internal Switch condition (red) during test blocks. Vertical lines are the onset and offset of the stimulus. Marks below show the channels and time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

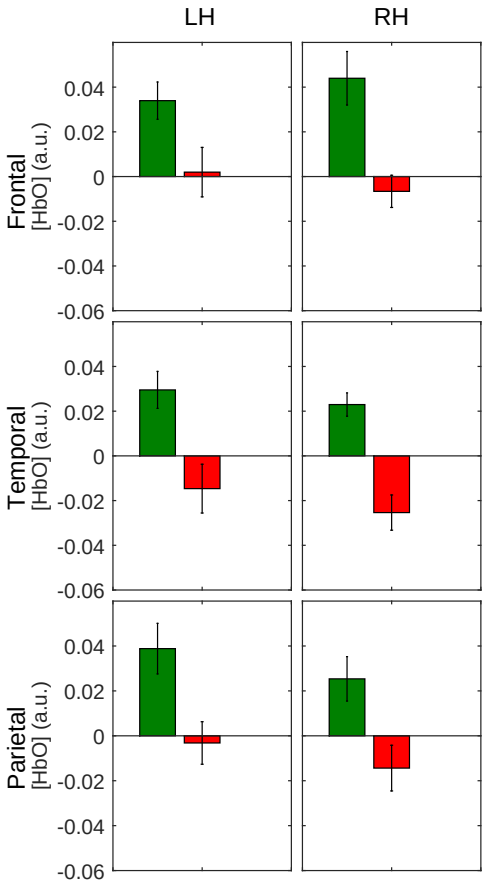


FIGURE 4.3: Mean activation analysis for Experiment 4 using HbO_2. In green activation during the Edge Switch condition, and in red activation during Internal Switch condition. Error bars represent standard errors

4.2.7 Discussions

Experiment 4 asked whether neonates better encode positional information for the items at the edges of a sequence compared to items in the middle. The results revealed that after being familiarized with a six-syllabic sequence, neonates showed a larger increase in HbO_2 during test blocks in which the two edge syllables switched position than when the two internal syllables switched positions. The neonates detected the change in the word when the edge syllables, but not when the internal syllables switched positions, indicating that the positional information of the edge syllables was better encoded than the positional information of the internal syllables. While previous results have demonstrated this phenomenon in adults (Endress and Mehler, 2009a; Endress, Nespor, and Mehler, 2009; Henson, 1998; Hurlstone, Hitch, and Baddeley, 2014) and in older infants (Benavides-Varela and Mehler, 2015), the current findings demonstrate that the constraints on sequential positional encoding are present from birth, before infants have extensive experience processing sequential information. This suggests that the edge bias is an inherent constraint in sequential processing that influences

how neonates process linguistic information.

The differential response between the Edge Switch and Internal Switch conditions emerged bilaterally on frontal and temporal areas. Even if Hb responses were noisier, we still found a similar pattern. This pattern of activation has been observed in neonate novelty detection experiments (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012; Nakano et al., 2009), as well as in the previous experiments reported in this thesis. As I have discussed for Experiments 1 and 2, frontal regions are involved in novelty detection across a range of tasks (Mahmoudzadeh et al., 2013; Nakano et al., 2009; Benavides-Varela et al., 2012) and the sensitivity of both temporal regions in speech processing has been evidenced in young infants and neonates (Minagawa-Kawai et al., 2011; Peña et al., 2003; Dehaene-Lambertz, Dehaene, and Hertz-Pannier, 2002). The broad, differential response across multiple channels is common in newborn fNIRS experiments, and we note that while we can conclude that there are clear differences between the conditions, conclusions regarding specific regions of activation must be treated with caution.

We explained our results on the absolute position of the switched syllables in the sequence, nevertheless an alternative explanation exists. Infants could have not noticed the change in the Internal Switch condition, because the distance between swapped syllables was shorter —in the Internal Switch syllable *i* and *i+1* were switched, whereas in the Edge Switch the exchanged syllables were *i* and *i+5*. Based on previous work with infants (Benavides-Varela and Mehler, 2015) and adults (Gupta et al., 2005; Henson, 1998) that do control for the distance between the switched items and found the same effect; we favour the hypothesis that also neonates notice the change because edges are better encoded. Furthermore, from a theoretical perspective it seems to us unprovable an interpretation based on the distance between switched elements: if infants do not form an abstract representation of position, but instead they encode association between the syllables at a local level (like transitional probabilities), we would expect the opposite pattern of results, meaning that changes involving items in closer positions would be easier to notice that changes regarding items in more distant positions. Said this, it remains an open question what neonates are actually encoding.

In sum, Experiment 4 revealed that even from birth, the encoding of syllabic positional information in a sequence depends of the position in the sequence. Edge positions seem to be more accurately encoded than internal positions. But what are the implications for actually processing language? Language requires encoding multiple hierarchical elements from a sequential stream —syllables combine to form words, words combine to form phrases, and phrases combine to form sentences. The positions of the items in each sequence are crucial to retain the meanings across each of these levels. How can the positional information across an entire sentence structure be encoded if the positions of internal elements are poorly encoded? One possibility is that prosodic segmentation cues (e.g., pauses or pitch contours) act to break up otherwise continuous speech and provide cues

to word edges and phrasal boundaries with sequential processing constraints operating across each segment. We address this question in Experiment 5.

4.3 Subtle pauses affect the encoding of a sequence of syllables. Experiment 5

Previous research suggests that prosodic cues can, in fact, signal the hierarchical constituent structure of language (Hawthorne and Gerken, 2014; Langus et al., 2012; Nespor and Vogel, 2007). For example, both adults and infants expect that words be contained within prosodic constituents delimited by boundaries and infants can better segment novel words from the edge of a sentence or phrase than from the middle (Johnson and Jusczyk, 2001; Langus et al., 2012; Seidl and Johnson, 2006; Shukla, White, and Aslin, 2011). Peña and colleagues (Peña et al., 2002) demonstrated that adults were unable to learn an AxC pattern of words from continuous speech but succeeded when a consciously imperceptible 25 ms pause was inserted between the words. These findings show that older infants and adults can use prosodic cues to segment continuous speech into discreet elements. In Experiment 5, we examined whether newborns use segmentation cues to break up a continuous sequence into smaller, discreet elements, with sequential processing constraints operating across each of the subcomponents.

As a cue for segmentation, we inserted a 25 ms silent pause in the middle of the sequence, between the third and fourth syllables (see Table 4.3). If this cue facilitates segmentation, the six-syllabic sequence would be divided into two shorter 3-syllabic sequences. The 3rd and 4th syllables —which were previously internal to the long sequence— would now be the final syllable of the first segment and the first syllable of the second segment, both in edge positions. If sequential processing constraints operate across each segment, and the 3rd and 4th syllables are now encoded as edges, the neonates would detect the change in the sequence when those syllables switch positions.

We tested an additional 16 neonates using an fNIRS testing protocol nearly identical to the one used in Experiment 4 (see Figure 4.1), with the only difference being the addition of the 25 ms pause during the familiarization and test sequences. The six-syllabic familiarization sequence contained a 25 ms silent pause between the third and fourth syllables (e.g., simebu-talɛfo, see Table 4.3). In the test blocks the same sequence (including the 25 ms pause) was presented, but the third and fourth syllables had switched positions (e.g., simeta-bulefo). The only difference between this condition (Pause Switch condition) and the Internal Switch condition of Experiment 4 was the addition of the 25 ms of silence between the two middle syllables. If neonates detected the syllable switch, the

HbO_2 would increase during test blocks (as in the Edge Switch condition of Experiment 4); if they did not detect the syllable switch, even with the addition of the segmentation cue, the HbO_2 would not increase during the test block (as in the Internal Switch condition of Experiment 4).

4.3.1 Participants

All participants were full-term neonates born to Italian-speaking mothers, with Apgar score ≥ 7 in the first minute and ≥ 8 in the fifth minute, diameter of head ≥ 33.0 cm, and no cefalhematomas. The Pause Switch Condition included 16 neonates (10 females; mean age 2.75 days, range 1–5 days; mean gestational age 39.0 weeks, SD = 1.0 week; mean weight 3.267 kg, SD = 0.361 kg), who provided data without motion artifacts from at least two of the six of the test blocks. Additional infants were excluded due to motion artifacts (n = 8) and failure to acquire a good signal due to thick hair (n = 4). All newborns were recruited from the nursery at Hospital, Azienda Ospedaliera Santa Maria della Misericordia, in Udine, Italy. Parents provided informed consent. The Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

4.3.2 Stimuli

The stimuli of the Pause Switch condition were almost identical to the stimuli of the Internal Switch condition, with the only difference that a 25 ms pause was inserted between the third and fourth syllables of the familiarization and test sequences (see Table 4.3). Stimuli were again synthesise using MBROLA (Dutoit et al., 1996). We included the pause in both familiarization and test sequences to ensure that infants responded to the switch in syllables and not to a change in the presence of a pause.

4.3.3 Procedure

Idem than for Experiment 4. See 4.2.3

4.3.4 Apparatus and data acquisition

Idem than for Experiment 1. See 2.3.4

4.3.5 Data Analysis

The analysis was done in analogue way than for Experiment 4.

Pre-processing. The pre-processing was identical than for Experiment 4 (see 4.2.5).

Pause Switch	
Fam	**Test**
simebu_talɛfo	simeta_bulɛfo
simeta_bulɛfo	simebu_talɛfo
nɛkali_sorevu	nɛkaso_lirevu
nɛkaso_lirevu	nɛkali_sorevu
gamɛzi_bekotu	gamɛbe_zikotu
gamɛbe_zikotu	gamɛzi_bekotu
neloki_subɛma	nelosu_kibɛma
nelosu_kibɛma	neloki_subɛma
ponive_lagusɛ	ponilavegusɛ
ponila_vegusɛ	ponivelagusɛ
boketa_sɛluma	bokesɛ_taluma
bokesɛ_taluma	boketa_sɛluma

TABLE 4.3: Stimuli for Experiment 5. A 25 ms pause was inserted between the third and four syllables. The test word was created exchanging the intermediate syllables. Different pairs of familiarization and test words were used. Infants were randomly assigned a pair of words.

HRFs calculation. The calculation of the HRF during the pause Switch condition was done as described for the Edge and the Internal Switch Conditions (see 4.2.5).

Data rejection. A HRF for a block and channel was excluded if during that period we detected either saturation, low signal to noise ratio, or not properly corrected motion artifacts. Considering all the infants tested in the Pause Switch condition, 48% of the total HRFs were automatically rejected and 9% manually rejected. Infants were included in the analysis only if they contributed at least with two good test blocks per condition. Each infant contributed an average of 3.25 test blocks (SD = 1.24).

Statistical Analysis. We performed a Cluster Based Permutation Analysis and a Mean Activation Analysis as previously. The Pause Switch condition was compared with the Internal Switch condition of Experiment 4.

4.3.6 Results

Results from both analysis are presented below. Figures and Tables for results based on Hb are presented in the Appendix C.

Cluster Based Permutation Analysis. We found a greater hemodynamic response for the Pause Switch condition than for the Internal Switch condition, mainly in fronto-temporal areas. The analysis revealed a significant cluster for HbO_2 in the

HbO_2	SumSq	DF	MeanSq	F	P
Condition	0.01550	1	0.01550	6.53951	0.02190
Error	0.03555	15	0.00237		
Hemisphere	0.00000	1	0.00000	0.00266	0.95950
Condition: Hemisphere	0.00275	1	0.00275	5.14905	0.03840
Error (Hemisphere)	0.00802	15	0.00053		
Region	0.00160	2	0.00080	1.79690	0.18320
Condition: Region	0.00181	2	0.00091	2.03886	0.14780
Error (Region)	0.01334	30	0.00044		
Hemisphere: Region	0.00117	2	0.00058	1.28723	0.29080
Condition: Hemisphere: Region	0.00091	2	0.00045	0.99936	0.38000
Error (Hemisphere: Region)	0.01360	30	0.00045		

TABLE 4.4: Statistical analysis for Experiment 5. 3-ways ANOVA on the mean $[HbO_2]$ as dependent variables. The hemisphere (left/ right) and the region (frontal/ temporal/ parietal) are within subject factors, and the condition (Edge/ Internal) is the between subjects factor

left hemisphere including channels 6, 7, 8 and 9, within the time window [8s - 16s] from the onset ($P_{HbO_2}cluster_1 < 0.05$); and one cluster in the right hemisphere with channels 13, 14, 15, 16 and 18 during the period [7s - 12s] ($P_{HbO_2}cluster_2 < 0.05$) (see Figure 4.4). For Hb we did not find significant results (see Figure C.3). We also compared the Pause Switch condition against the Edge Switch condition but we did not find significant results neither with HbO_2, nor with Hb (P>0.05).

Mean Activation Analysis. Using HbO_2 as dependent variable we found a main effect of condition (Edge / Pause) (F(1,15) = 6.5395, P < 0.05), and a significant interaction Condition: Hemisphere (F(1,15) = 5.1491, P < 0.05), and not other main effects or interactions (P > 0.05) and (see Figure 4.5 and Table 4.4). Post-hoc multiple comparisons Turker-Kramer corrected, revealed that the interaction was due to a significant difference between conditions in the right hemisphere (P < 0.01) but not in left hemisphere (P > 0.05). As with the cluster based permutation analysis we did not find significant results for Hb (see Figure C.4, and Table C.2). By comparing the Pause Switch condition against the Edge Switch condition we did not find significant results (P > 0.05).

4.3.7 Discussions

Experiment 5 revealed that neonates detect a positional switch between two internal syllables in a sequence if those syllables are separated by a 25 ms pause. We asked if newborns use prosodic boundaries to break up continuous speech into smaller segments, with sequential processing constraints operating across each of the segments. While neonates failed to notice a positional switch between the middle syllables of six-syllabic sequence in Experiment 4, they did detect the position switch when a subtle 25 ms pause was inserted between the middle syllables, suggesting that the pause facilitated encoding of the positional information

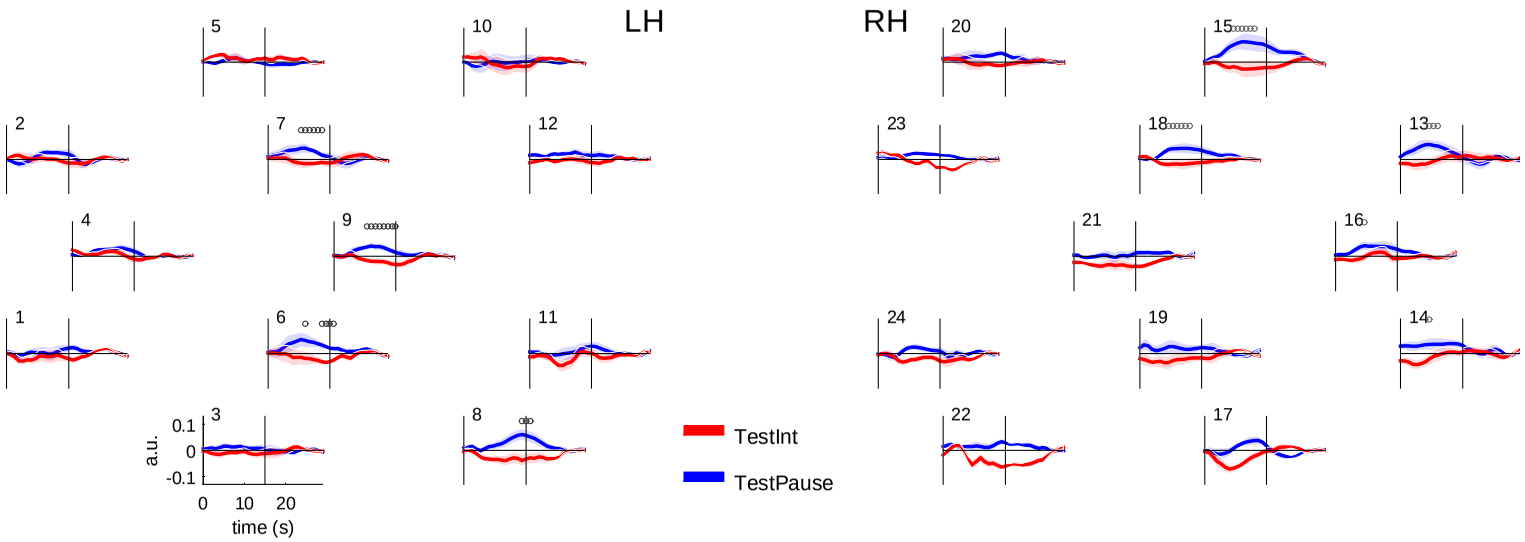


FIGURE 4.4: Cluster based permutation analysis for Experiment 5 using HbO_2. HRFs for the Pause Switch condition (blue) and Internal Switch condition (red) during test blocks. Marks below show the channels and time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

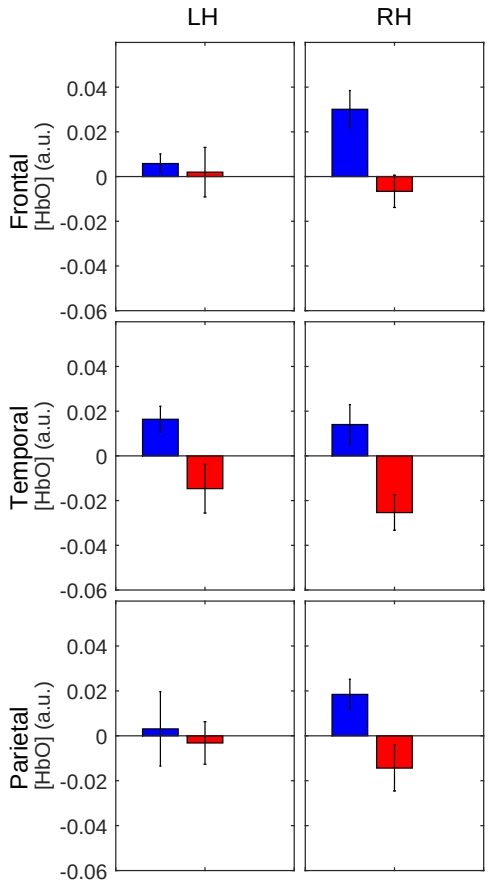


FIGURE 4.5: Mean activation analysis for Experiment 4 using HbO_2. In blue activation during Pause Switch condition, and in red activation during Internal Switch condition. Error bars represent standard errors.

of the otherwise internal syllables. We propose that the 25 ms pause in Experiment 5 segments the six-syllabic sequence into two three-syllabic sequences, and that the sequential processing constraints operate across each sequence individually. With the pause, the 3rd and 4th syllables were encoded as edges —right and left, respectively— of each smaller segment, and the neonates detected the change when those syllables switched positions.

As in Experiment 4 the differential pattern of increase in HbO_2 between the Internal Switch condition and the Pause Switch condition emerged bilaterally, and in regions comprising mostly temporal and frontal areas. The pattern of results between the Pause Switch condition and the Edge Switch condition were similar, with the cluster analyses revealing a majority of significant channels overlapping between the two conditions. The results from the mean activation analysis support this conclusion showing increases in HbO_2 for the Pause Switch condition respect to the Internal Switch condition. This analysis also revealed the effect was stronger on the right than on the left hemisphere. This is consistent with previous experiments reporting stronger activation toward novel stimulus in frontal right

then in frontal left regions (Nakano et al., 2009). Unlike in Experiment 4 we did not find significant results for Hb. A possible explanation is that the change was indeed harder to notice than in the Edge Switch condition and Hb did not provide significant results because it has a lower signal to noise ratio.

Experiment 5 also provides additional evidence that the results obtained in Experiment 4 are due to the more precise encoding of edges, rather than more low-level explanations, such as the number of syllables between the switched elements. Our converging evidence demonstrates that, with the same syllabic sequences, neonates detect a switch between two adjacent internal syllables only when provided with a cue indicating an edge between those syllables. Inserting an edge between the internal syllables enhances their encoding indicating that specifically the edges of sequences are more precisely encoded.

A more thorough discussion may ask if it is really an edges what the 25 ms pause generates, meaning by an edge that infants form two representation of three-syllable words instead of a unique representation of a six-syllabic sequence. The role of the pause may involve a switch of attention toward syllables close to it, which translates in its better encoding. This is a valid discussion that we are not able to fully answer base on the results. However, an attentional switch implies a disruption in the processing, thus it involves producing a break. Regardless if one or two short term memory representations are generated the general conclusion remains the same: there are positions in a sequences —the edges— that are better encoded since birth, and subtle pauses affect how sequential information is encoded by enhancing the encoding of syllables close to it.

4.4 Chapter Discussions

In two neuroimaging experiments we demonstrated that neonates encode sequential information from multisyllabic sequences and better encode positional information from sequence edges. From birth, infants can also utilize very brief pauses to segment longer syllabic sequences into smaller sequences, with these positional constraints operating across both smaller sequences. These results, demonstrating that sequential processing biases constrain how linguistic stimuli are encoded from birth, are crucial for our understanding of the mechanisms underlying language acquisition in two key ways.

First, Experiment 4 demonstrated that humans are born with processing constraints that privilege sequence edges; neonates are more sensitive to positional changes of syllables at the edges of sequences than to the internal syllables. While previous work had demonstrated an edge processing bias in adults and older infants (Benavides-Varela and Mehler, 2015; Henson, 1998; Hurlstone, Hitch, and Baddeley, 2014), it remained unclear whether this constraint was the result of a fundamental constraint on sequential processing or a consequence of exposure to sequential processing. For example, language highlights edges by cues such

as determiners, morphological regularities, and stress and exposure to these cues might act as to make edges more salient. We demonstrated that even from birth, there is an enhanced encoding of sequence edges, indicating that this constraint is an inherent signature of sequential processing. The enhanced encoding of edges constrains linguistic regularities to appear at the salient components of sequences rather than experience with edge-based linguistic regularities causing constraints in general sequential processing.

Second, Experiment 5 revealed that with a segmentation cue in the middle of the sequence, neonates detected the positional switch of the previously internal syllables. From birth, infants use prosodic cues to segment long sequences into smaller sequences, indicating that sequential processing constraints can operate across these multiple segments in parallel. This ability is fundamentally important for acquiring language since language requires encoding multiple hierarchical levels (e.g., encoding the order of syllables in words and the order of words in sentences) from a single sequential stream. Our results suggest that the foundational mechanisms for tracking across different hierarchical levels are evident from birth. Prosodic cues have been found to mark boundaries across different hierarchical levels (Langus et al., 2012; Nespor and Vogel, 2007), and we demonstrate that newborns are capable of using subtle cues to segment and encode longer sequences. Our results suggest that newborns can use prosodic boundaries embedded in the speech signal to track the edge of units in the sequential stream of speech. This mechanism may be a possible explanation for how infants begin to encode language across multiple hierarchies.

These findings also give new insight regarding how language universals may emerge from the inherent processing constraints. First, across languages, short words are much more frequent than longer words (Sigurd, Eeg-Olofsson, and Weijer, 2004), which may be the result of less efficient encoding of the internal components of longer words. Second, across languages, components at the edges of linguistic units (e.g., words) are generally crucial to the structure of language (e.g., determiners, bound morphemes) (Greenberg, 1957). Our results demonstrating an enhanced encoding of edge components from birth raise the possibility that linguistic regularities occur at the edges because of inherent human processing constraints. These components may occur almost universally at the edges because regularities are more easily learned at the edges. Third, linguistic stress is universally encoded with respect to boundaries; for example, lexical stress cues boundaries between words (Greenberg, 1957; Hayes, 1995). Fourth, the finding that neonates use a subtle prosodic cue to segment long sequences adds to our finding in Experiment 2 to the role of prosody in early language acquisition. Prosodic cues may facilitate language acquisition even during the first days of life.

These findings raise questions about the specificity of these processing constraints. First, while we focus on syllabic sequences, these constraints on sequential processing potentially extend to non-linguistic sequences as well. The edge bias in sequential processing is robust in adults, emerging across a range of tasks and sensory domains (Endress and Mehler, 2009a; Endress, Nespor, and Mehler, 2009; Fournier et al., 2014; Gupta et al., 2005; Gupta, 2003; Hurlstone, Hitch, and Baddeley, 2014; Murdock, 1962). If the sequential encoding mechanism is general, then similar results should emerge with non-linguistic stimuli as well, though potentially with domain-specific segmentation cues. Since the current experiments only used sequences of syllables, the generality of sequential processing at birth and the generality of the role of segmentation cues in sequential processing remains a question for future research. Second, natural language contains cues (e.g., pauses, stress markers, prosodic contours, determiners, transitional probabilities) that mark different hierarchical boundaries (Gervain and Werker, 2013; Greenberg, 1957; Hawthorne and Gerken, 2014; Hirsh-Pasek et al., 1987; Hochmann, 2013; Halle and Vergnaud, 1987; Nelson et al., 1989; Saffran, Newport, and Aslin, 1996) and each of these boundaries could signal an edge. Since we focused our experiment specifically on the brief pause as a boundary cue, it remains an open question whether other prosodic boundaries can similarly facilitate sequence segmentation and whether they do so from birth.

To conclude, our findings advance the understanding of language acquisition by addressing how neonates encode multisyllabic sequences. We revealed that not only do newborn infants encode positional information from multisyllabic sequences, but that they do so in a fundamentally constrained way, encoding positional information from the edges of sequences much more precisely than from internal positions. This discovery reveals that even at birth, language processing is constrained, with some information being more precisely encoded. Moreover, we revealed that neonates utilize very subtle cues to segment a continuous sequence into smaller sequences, with the sequential processing constraint operating across each subcomponent. In sum, these findings reveal that humans are born with specific constraints on how language is processed and the ability to use subtle cues in the speech signal to segment the sequential streams of syllables that comprise language. The hierarchical organization of syllabic sequences is a fundamental signature of language and our results suggest that the foundational mechanisms to track this organization are present from birth.

Chapter 5

How is a continuous flow of stimuli encoded?

5.1 The encoding of position, rules and prosody: How do they interact?

One of the big question in language acquisition is the bootstrapping problem. How do infants learn the constituent of language? and moreover how do they acquire its organization? Language is by nature sequential, which leads to the segmentation problem I discussed in Chapter 2. Units have to be identified, both prosodic units, like utterances or intonational phrases, as sentences and its constituents, the words. But the issue does not restrict to segmentation, the structure underling language has to be uncovered.

From the generative grammar approach sentences' structure is described using syntactic trees (Chomsky, 1957), denoting the hierarchical organization of language. Under this view hierarchy is the outcome of multiple "merge" operations, whereas the linear organization of language is consequence of a constrain of the sensory-motor system to present the phonemes one at the time (Berwick et al., 2013). This view is questioned by some authors that consider that language is indeed characterized by its sequential structure, and that it can be learned without appealing to a hierarchical representation, which is computationally too demanding (see for example Frank, Bod, and Christiansen, 2012; Christiansen, Conway, and Onnis, 2012). The two approaches imply very different mechanism for how the rules underlying language structure are discovered. Whereas the former has to rely on some type of abstract representation (not based on the surface of the stimuli), the latter is based on simple associations between items.

Leaving aside the semantic aspect of the problem, our capacity to segment the speech, extract regularities and eventually form rules, depends on the computations operating over the signal —statistical learning, perceptual biases— and of our capacity to encode the information —memory constraints. A basic question emerges: how is the information encoded when units in continuous speech are identified? One possibility is that we rely on the co-occurrence of the constituent to identify the units, which would be coherent with an intrinsic sequential nature

of language. Another option is that a more abstract representations of the units is generated, in line with a hierarchical nature. For example if we listen something like *the boy plays, my cat runs, a bird sings*, the former options implies we create simple associations between the items, meaning that we encode that *plays* appears after *boy*, which appears after *the*. Under the later view we should encode that *the*, *my* and *a* have the same role in the sentences, which is different from the role of *boy*, *cat* and *bird*, and from *plays*, *runs* and *sings*. With the previous example we wanted to make noticeable how the extraction of structure from serially presented stimuli may depend, or be reflected, in how position is encoded.

We extensively discussed serial position encoding in Chapter 4, but it is worth it to go through some relevant aspects in more detail. Short term memory studies suggest that a sequence of items is not —or not only— represented based on the associations between the item, but that the position of the items in a sequence is encoded relative to the edges (Ebbinghaus, 1885; Hurlstone, Hitch, and Baddeley, 2014; Endress, Nespor, and Mehler, 2009; Henson, 1998). This hypothesis derives from the observation of various effects: 1. Items in the edges are remembered better —primacy and recency effects. 2. Transpositions (the exchange of two elements) are more frequent in intermediate positions. 3. Items occupying the same position, but belonging to different lists are easily confused (Henson et al., 1996). 4. If verbal sequences are grouped with temporal pauses the accuracy of recall increases, but it also increases the number of transpositions between groups, in particular between items that share the same position in the group (Henson et al., 1996; Henson, 1999).

An abstract encoding of position seems particularly suitable for a system like language, in which positional information is fundamental, and at the same time items can be exchanged as long as they belong to the same class. But, language flows continuously along time, thus, where are the edges? thereby leading back to the segmentation problem. For an abstract serial encoding to be effective for linguistic representations, we should be able to create such representations even when stimuli are not presented in isolation.

In this chapter we investigate if positional information relative to edges is encoded either when information is presented continuously, and when subtle pauses signalled the edges. Our previous experiment (Experiment 5) shows that subtle pauses effect how neonates encode a six-syllabic word. Moreover, previous studies show that subtle pauses affect adults capacity to generalize structures. Adults succeed to extract words marked by long distant dependencies of the form AxC, where A predicts C, but they are able to generalize the rule only when 25 ms pauses are inserted between the words (Peña et al., 2002). Furthermore, an increase in the power spectrum of the EEG signal at the frequency of the words is stronger when pauses separate the words (Buiatti, Peña, and Dehaene-Lambertz, 2009), which was also observed in 8-month-old infants (Kabdebon et al., 2015). We were interested on knowing if an abstract encoding of position is created when

the structure of a continuous flow of stimuli is learned, and if the formation of this more abstract representation depends of the presence of other cues besides distributional information.

From our perspective, it is unlikely that position is not tracked and that everything relies on simple associations. For example, long distance dependencies are very common in language, but tracking associations between distant items seems computational and memory too demanding. On the contrary, encoding position appears as a more efficient mechanism. Moreover, the encoding of positions inside units can be considered as a first requirement for building a hierarchy. A second requirement would be to generate categories between items —exchangeable items. An finally, position should be tracked over different levels, so that an item of one level is an unit of a lower level. In the current experiments we do not explore this last point, but we hypothesize that, because prosody has various components associated with hierarchical levels (Nespor and Vogel, 2007), it is a good candidate to mark parallel edges and levels of processing. In the current experiments we investigate the role of a single prosodic level (subtle pauses), which we think may contribute to our understanding of the representation that is generated of language structure.

There is a last aspect of our hypothesis that I would like to remark. An abstract encoding based on the position relative to the edges implies associating items with its location, which provides another dimension across which items can be linked and generalizations made. This introduces a new perspective in the relation between statistical learning and rule learning, which have traditionally been considered as an independent learning mechanism. Rule learning is the generalization of a pattern to new instances. For example we know that 7 month-old can discriminate the rule ABB from ABA after listening to segmented words (G. F. Marcus, S. Vijayan, S. Bandi Rao, 1999). The interpretation is that infants create an abstract representation of this pattern that afterwards generalize to unfamiliar items [1]. Another example of rule learning in this context is Peña et al. (Peña et al., 2002) experiment. In both examples if positional information is encoded a new feature —position— is associated to each of the items, and generalization could be the conservation of position. Of course we are referring to a very particular type of generalization, generalization in more broad sense implies much more and has nothing to do with our studies (see for example Gerken, 2010; Aslin and Newport, 2012).

[1] Another study shows that neonates recognize ABB form ABC but not from ABA from ABC (Gervain et al., 2008a), leading to think that maybe infants are sensitive the repetitions due to its perceptual saliency (Endress and Mehler, 2009b)

5.2 Adults segmenting linguistic auditory stimuli. Experiment 6

We hypothesised that once the edges in a continuous stream of syllables are found, together with the emergence of edges another way of encoding the information arises: each syllable can be associated to a position relative to the edges, and potentially this could become more crucial in the representation of the segmented chunks than the co-occurrence of syllables. If this is the case, it has important consequence on how we uncover the rules governing language, a system for which position is fundamental. Coming back to our original example, if we listen *the boy plays, my cat runs, a bird sings*, based on position we would accept as a possible sentence *the cat plays* even if we have not heard it before. In addition to the distribution of the syllables, natural language has a rich prosody, which can also signal units. Thus different panoramas are possible. One possibility is that an abstract encoding is generated only when the sentences are presented in isolation, or in practical terms, when prosodic cues signal the edges. Another option is that an abstract encoding is generated independently of segmentation cues. A final possibility is that a more abstract representation is not generated at all.

In a first study, we decided to test this hypothesis in adults in a very simple behavioural experiment. We asked participants to find the words from a continuous stream of syllables, as it has been classically done in segmentation experiments. In order to test our hypothesis, we did not need complicate dependencies between syllables, rather we wanted the task to be as easy as possible, thus we used a familiarization stream with the same structure of the one used by Saffran and colleagues in their first experiment (Saffran, Aslin, and Newport, 1996). The main difference we introduced with respect to their experiment, was a new word category during the test phase. The idea behind this new word category is that are sequence that have never appeared in the stream, but in which the position of the syllables is the same than in real words. Specifically, we built these sequences by exchanging the first syllables of two words, thus we called them Switch-words. Note the following (see Figure 5.1): for Words the position of the syllables and the frequency of appearances of bi-grams or tri-grams is perfects; for Part-words, bi-grams and tri-grams, even if less frequently than for Words, appear in the familiarization, however the position of the syllables relative to edges is not correct; finally Switch-words are conformed by syllables that never appear together, but each syllable preserves its position relative to the edges.

Participants first heard a familiarization phase, in which the stream was either continuous or in which the words were separated by 25 ms pauses. After the familiarization phase participants underwent force choice trials in which the three categories were compared between each other. We hypothesized that if after finding the edges adults start encoding the position of the syllables relative to the

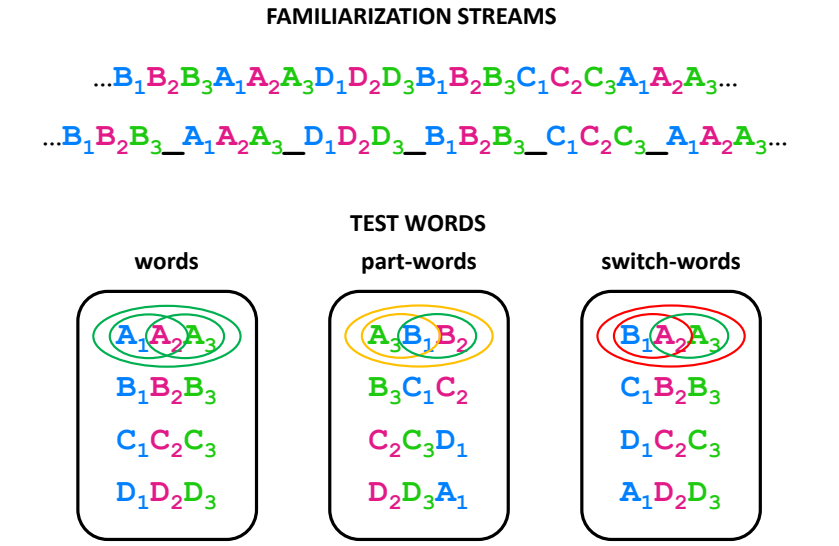


FIGURE 5.1: Scheme of the stimuli for the experiments of Chapter 5. A continuous stream and stream with brief pauses between the words were used. The Switch-words are conformed by the first syllable of one word and the last two syllables of another word. Part-words are sequences that appear in the stream even if with less frequency that Words (TPs > 0), but in which the syllables do not occupy the same position than in Words. Switch-words are sequences that never appear in the stream (TP between the first and second syllable = 0), but in which each syllables appears in the same position than in a Word.

edges, they should prefer Switch-words toward Part-words. If the abstract encoding occurs only when the edges are marked by a segmentation cue different from the co-occurrence of syllables, this preference should appear only for participants that were familiarized with the stream with pauses.

Finding that participants familiarized with the segmented stream prefer Switch-words over Part-words, would go in line with Peña and colleagues study (Peña et al., 2002), in which they showed that adults fail to generalized AxC rules when speech is continuous, but that they succeed when it is already segmented by 25 ms pauses. The authors interpreted their results as consequence of rule and statistical learning being two different and independent learning mechanism: in order to leave place to rule learning to occur, the units have to be marked by a cue that is not the dependencies between the syllables, otherwise statistical learning prevails and words are segmented but the rule is not generalized. Nevertheless, we think that this is not the only possible explanation. The pauses make the segmentation task easier, thus affect the level of performance, meaning by performance how well the words are extracted and represented. In order to disentangle if effects are due to the level of segmentation achieved or to the type of cue marking the edges during the familiarization, we checked if participants that achieved better segmentation, independently of the presence of pauses, prefer Switch-words over

Part-words.

5.2.1 Participants

Participants were 55 native Italian speakers that did not report any auditory or language related disorder. 23 infants were included in the continuous condition (7 males, mean age 24 years old, range [18, 31]), and 22 in Pause condition (9 males, mean age 24 years old, range [20, 29]). All participants received a monetary compensation.

5.2.2 Stimuli

The auditory stimuli consisted in a Familiarization stream that differed between the Continuous and the Pause condition, and three types of Test items that were identical between conditions. The general procedure used to build the familiarization stream was the same used in Experiment 1 (see 2.3.2). Specifically for this study the average transitional probability between words was 0.3323 (SD = 0.0108, range [0.3165 0.3457]); each Word appeared 80 times and the Part-Words 25 to 28 times. The syllables used had all a consonant-vowel structure, there were not repeated syllables, and no Italian words were present in the stimuli. Stimuli were synthesized using the it4 Italian female voice of the MBROLA diphone database (Dutoit et al., 1996), with phoneme duration of 150 ms and a constant pitch of 200Hz. In the Continuous condition sequences were continuous with no pauses between syllables. In the Pause condition 25 ms pauses were inserted between the words. Note that the pause disrupts the co-articulation.

For the Test items we used four examples for each type of test (see Table 5.1): (1) type Words (W condition), the four words used to synthesize the stream ; (2) type Part-words (P condition), two sequences of the form $A_3B_1B_2$ and two of the form $A_2A_3B_1$; (3) type Switch-words (S condition), built by replacing the first syllable of one word by the first syllable of another word.

In order to avoid that results were driven by specific properties of the stimuli, two different Familiarization streams and sets of Test items were created by shifting of one position how the syllables were taken to construct the words (see Table 5.1). Observed that half of the Part-words used of Stream A are Words of Stream B and vice-versa. Participants were randomly assigned to one or the other group.

5.2.3 Procedure

The continuous/pauses condition was between subjects; whereas the types of word comparisons (words/ part-words/ switch-words) were compared within subjects.

Before starting the experiment, participants overcame a training phase to get use to the response keys. During the training participants heard the syllable *ba* ten

Stream	W	P	S
A	baʃetsi	ʃetsife	doʃetsi
	dokave	gadoka	nikave
	nitoga	venito	fetoga
	fepiʤo	piʤoba	bapiʤo
B	ʃetsido	tsidopi	katsido
	kaveni	fekave	toveni
	togafe	nitoga	pigafe
	piʤoba	ʤobaʃe	ʃeʤoba

TABLE 5.1: Stimuli for Experiment 6. Words, Part-words and Switch-words used in the experiment. Two streams and sets of words were created and participants were randomly assigned one.

times and afterwards completed six test trials. In each test trial two syllables were presented—one of which was *ba*— and subjects had to decide which one they had heard before. If they thought it was the first, they were instructed to press the left button, whereas if they thought it was the second, the right button. Participants repeated the training till having 100% of correct responses.

Once they completed the training, the experiment started. Participants were told they will listen a sequence of syllables for some minutes and that they should try to identify the words conforming it. The total duration of the familiarization was 4 minutes for the Continuous condition and 4 minutes 8 seconds for the Pause condition. After the familiarization task participants performed a force choice recognition task. They heard to two three-syllables sequences and were asked to choose the one they though was more likely part of the stream. Three types of comparison were made: a) word vs. part-word b) word vs. switch-word c) part-word vs. switch word. The four words, part-words and switch-words were compared in 16 force choice trials per type.

5.2.4 Apparatus and data acquisition

The experiment was written in MATLAB, using the Psychophysics Toolbox extensions (Brainard, 1997), and run in a iMAC 10.1 computer. It took place in a silent booth. Audio stimuli were delivered via headphones.

5.2.5 Data Analysis

For each type of force choice trial we calculated the proportion of times each subject chose one type of test item. We first compared each test trial type against chance level (0.5) using multiple t-tests. We applied Bonferroni to correct for multiple comparisons. In order to evaluate the effect of pauses during the familiarization phase on the preference for words over part-words and over switch-words, we used a two way ANOVA. We included the familiarization phase condition (continuous/pause) as between subjects factor, and type of comparison (W vs. P

	WvsP		WvsS		PvsS	
	c	p	c	p	c	p
P	0.0000	0.0000	0.0000	0.0000	0.5620	0.0040
P corr.	0.0001	0.0000	0.0001	0.0000	4.4959	0.0290

TABLE 5.2: Statistical analysis for Experiment 6. t-test against chance level for the percentage of responses for each type of force choice trial and familiarization condition. Bonferroni correction was applied to correct for multiple comparisons.

/ W vs. S) as within subjects factor. The force choice trials part-words vs. switch-words were not included in this analysis, because this would be miss-leading —the comparison is transversal to the other comparisons involving words. We also ran an ANOVA including the stream used during the familiarization (A/B) as a between subjects factor, but because we did non find any effect or interaction ($P > 0.05$) we did not include it as a factor in the final analysis reported here.

To disentangle what was consequence of the level of performance and what to the presence of the pauses, we investigated which subjects where showing a stronger preference for switch-words vs. part-words relative to their preference for words over either part-words or switch-words. We plotted the proportion of answers in each force choice trial against each other in order to examine the distribution of their answers, and we checked if subjects with a stronger preference for Words over Part-words (better performers) also prefer Switch-words over Part-words.

5.2.6 Results

The proportion of answers for each type of words are presented in Figure 5.2. The t-tests against chance level confirmed that subjects preferred Words over Part-words and Switch-words in both conditions (see Table 5.2), which indicates that they extracted the words. On the contrary we found a significant preference for Switch-words over Part-words only for the group familiarized with the stream with pauses. Moreover, the 2-ways ANOVA revealed an effect of familiarization phase (continuous/pause) ($F(1,43) = 10.53$, $P < 0.01$), of force trial comparison (W vs. P / W vs. S) ($F(1,43) = 7.55$, $P < 0.01$), and a significant interaction ($F(1,43) = 4.67$, $P < 0.05$) (see Table 5.3). Post-hoc multiple comparison analysis, using Tukey-Kramer correction, revealed that the interaction was due to a higher preference for Words over Part-words than over Switch-words, only in the pause group ($P < 0.05$), and to a difference between the pause and continuous condition only for the Words vs. Part-words trials (see Table 5.4).

Afterwards, we investigated the pattern of responses of individual subjects by plotting the proportion of answers in each force choice trial against each other (see Figure 5.10). As expected, subjects with a preference for Words towards Part-words also prefer Words towards Switch-words —data points in Figure 5.3 appear mainly on the upper right quadrant. Interestingly, data suggests that only

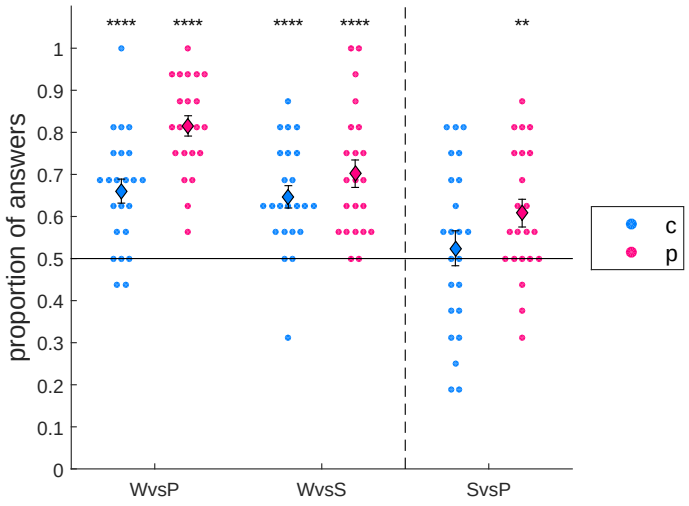


FIGURE 5.2: Answers for each type of force choice test trial after participants were familiarized with a stream of syllables. Each dot represents the proportion of answers for one subject in a specific comparison. Results for the group familiarized with the continuous stream is plotted in blue, and with the stream with pauses in red.

	SumSq	DF	MeanSq	F	P
Intercept	44.84	1.0	44.84	1905.11	0.00000
Between groups	0.25	1.0	0.25	10.53	0.00230
Error	1.01	43.0	0.02		
Within groups	0.09	1.0	0.09	7.55	0.00870
Interaction	0.06	1.0	0.06	4.67	0.03630
Error	0.52	43.0	0.01		

TABLE 5.3: Statistical analysis for Experiment 6. 2-ways ANOVA on the percentage of responses as dependent variable. The between factors is the familiarization phase condition (continuous/ pauses) and the within factor the force choice trial type (W vs. P/ W vs. S)

for the group familiarized with the continuous stream, performance in the force choice trials Word vs. Switch-word is above chance even for subjects performing at chance in the force choice trial Word vs. Part-word. When the answers for the trials Words vs Part-words and Switch vs. Part-words were plotted against each other, we observed that for both groups, subjects showing a higher preference for Words towards Part-words, prefer Switch-words to Part-words —answers are negatively correlated (see Figure 5.3.b).

5.2.7 Discussions

Subjects had a clear preference for Words over Part-words and Switch-words as expected, meaning that they segmented the stream and extracted the words. Their preference was modulated by the presence or not of pauses segmenting the

	Difference	StdErr	P	CI lower	CI upper
c: WvsP - WvsS	0.01359	0.03236	0.67671	-0.05168	0.07890
p: WvsP - WvsS	0.11364	0.03309	0.00133	0.04690	0.18040
WvsP: c - p	-0.15501	0.03756	0.00017	-0.23076	-0.07930
WvsS: c - p	-0.05497	0.04188	0.19631	-0.13942	0.02950

TABLE 5.4: Statistical analysis for Experiment 6. Multiple comparison post-hoc analysis. Tukey-Kramer correction was used.

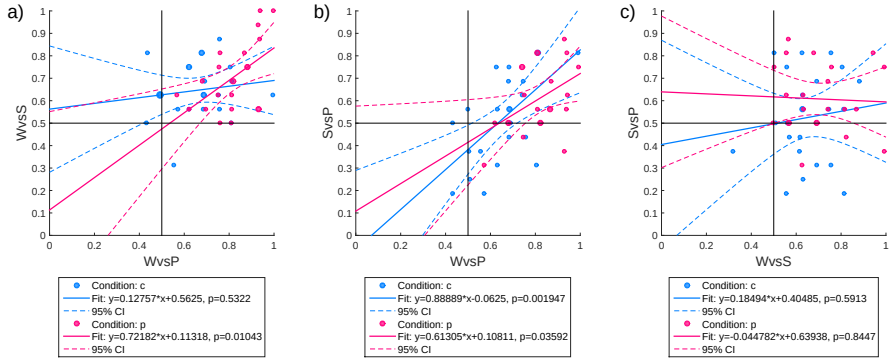


FIGURE 5.3: Individual answers for Experiment 6 for the three types of force choice trials plotted against each other. The size of each dot represents the number of subjects for that data point. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Answers for Words vs. Part-words force choice trials in the x axis, and for Words vs. Switch-words in the y axis. b) Answers Words vs. Part-words force choice trials in the x axis, and for Switch-words vs. Part-words in the y axis. c) Answers for Words vs. Switch-words force choice trials in the x axis, and Switch-words vs. Part-words in the y axis.

stream: participants familiarized with the stream with pauses had a higher preference for Words, which was also expected. What it is interesting in the results is that pauses did not simply increase the preference for Words over Part-words and Switch-words, but that there was an interaction: the preference for Words over Part-words increased, but not the preference over Switch-words. Moreover, participants familiarized with the continuous stream did not show any preference between Part-words and Switch-words, whereas participants that heard the stream with pauses chose the Switch-words to the Part-words.

Notwithstanding this last finding, the preference for Switch over Part-words was positively correlated with the preference for Words over Part-words in both the continuous and pauses conditions. This suggest that it is the level of performance in the segmentation task and not the type of cue present in the stream what affects subjects preference.

A consideration should be done regarding the subtle pauses in the familiarization stream. These pauses are not only silences of 25 ms, but they play the role of disrupting the co-articulation between the syllables. Anyhow, this does

not modified the conclusions. The pauses acts as a prosodic cues that help for the segmentation task.

Summarizing, our results suggest that once a continuous stream of syllables is segmented, independently if the edges are marked by transitional probabilities or by subtle pauses, the relative position of the syllables respect to edges starts to be encoded, and this information becomes more relevant for defining the identity of the segmented sequences than the co-occurrence of the syllables. This result has important consequences for language, in which positional information is fundamental.

5.3 Adults segmenting non-linguistic auditory stimuli. Experiment 7

In Experiment 6 we showed that when continuous speech is segmented, independently of how, the co-occurrence of syllables is not the only piece of information that is encoded, and that a more abstract encoding based on the position inside the chunks of syllables appears. This is potentially a good mechanism for encoding linguistic information, which begs the question: is it a general mechanism that works across domains or is it a language specific phenomenon? In order to answer this question we decided to repeat the same experiment, but using auditory non linguistic stimuli. We tested other two groups of subject, replacing the twelve syllables of the Experiment 6 by twelve pure tones. Considering that in a tone sequences there is not co-articulation to disrupt, and in order to make the pause more noticeable, we inserted longer longer pauses (50 ms).

If an abstract encoding is a general encoding mechanism, we expect to find similar results than in Experiment 6. Instead if it relies on language specific mechanism we should observe that subjects prefer Words over Part-words and Switch-words (they segment the stream); but they prefer Part-words over Switch-words, because Part-words are sequences that they have heard, whereas Switch-words are completely new to them.

5.3.1 Participants

Participants were 55 native Italian speakers that did not report any auditory or language related disorder. 23 participants were included in the continuous condition (10 females, mean age 24 years old, range [20, 30]), and 22 in the Pause condition (10 females, mean age 24 years old, range [19, 32]). All participants received a monetary compensation.

5.3.2 Stimuli

The familiarization streams and test items had identical distributional properties than in Experiment 6 (see 5.2.2), but the twelve syllables were replace by twelve

Stream	W	P	S
A	AFB	FBC	$F^\sharp FB$
	$F^\sharp A^\sharp D$	$D^\sharp F^\sharp A^\sharp$	$EA^\sharp D$
	$EGD^\sharp$	DEG	$CGD^\sharp$
	$CG^\sharp C^\sharp$	$G^\sharp C^\sharp A$	$F^\sharp G^\sharp C^\sharp$
B	$FBF^\sharp$	$BF^\sharp G^\sharp$	$A^\sharp BF^\sharp$
	$A^\sharp DE$	$CA^\sharp D$	GDE
	$GD^\sharp C$	$EGD^\sharp$	$G^\sharp D^\sharp C$
	$G^\sharp C^\sharp A$	$C^\sharp AF$	$GC^\sharp A$

TABLE 5.5: Stimuli for Experiment 7. Words, Part-words and Switch-words used in the experiment. Two streams and sets of words were created and participants were randomly assigned one. All tones belong to the fourth octave.

pure tones all belonging to the fourth octave (see Table 5.5). Each tone lasted 250 ms and was generated by a sin wave. The intensity was modulated by a Gaussian function with standard deviation 1/4 the duration of the tone. In the Continuous condition tones were concatenated with no pauses added between them. In the Pause condition 50 ms pauses were inserted between tones of different words. We decided to insert longer pause than in Experiment 6 to make them more noticeable, given that there is not something analogue to the co-articulation that gets disrupted in this case.

5.3.3 Procedure

The procedure was identical to Experiments 6 (see 5.2.3), but the syllables stream and test items were replaced by the tones stream and test items. During the training the syllables were also replace by tones.

5.3.4 Apparatus and data acquisition

Idem than for Experiment 6. See 5.2.4.

5.3.5 Data Analysis

Idem than for Experiment 6. See 5.2.5.

Also in this case we did non find any effect or interaction for the familiarization stream (A/B) ($P > 0.05$), thus we did not include it as a factor in the final analysis reported here.

5.3.6 Results

The t-tests against chance for each test trial (see Table 5.6) together with the 2-ways ANOVA (Tables 5.7), revealed that subjects did not really extract the structure in the stream of tones. In fact, after correcting for multiple comparison, subjects

	WvsP		WvsS		PvsS	
	c	p	c	p	c	p
P	0.0111	0.3922	0.0001	0.0138	0.4248	0.4440
P corr.	0.0890	3.1379	0.0011	0.1105	3.3984	3.5490

TABLE 5.6: Statistical analysis for Experiment 7. t-test against chance level for the percentage of responses for each type of force choice trial and familiarization condition. Bonferroni correction was applied to correct for multiple comparisons.

performed above chance only for the force choice trials Words vs. Switch-words in the continuous condition ($P < 0.05$) and marginally significant above chance in the pause condition ($P = 0.11$). Considering we do not have evidence of learning the tones sequences structure, exploring the patterns of answers for individual subjects seems pointless. However, data is presented in Figure 5.5 for descriptive purposes.

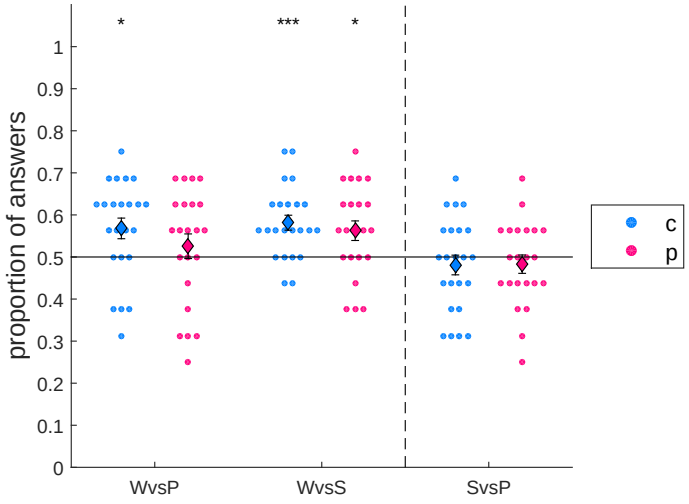


FIGURE 5.4: Answers for each type of force choice test trial after participants were familiarized with a stream of tones. Each dot represents the proportion of answers for one subject in a specific comparison. Results for the group familiarized with the continuous stream are plotted in blue, and with the stream with pauses in red.

5.3.7 Discussions

We did not find significant evidence indicating that adults were able to segment the stream of tones, clashing previous studies (Saffran et al., 1999; Abla, Katahira, and Okanoya, 2008). The tones sequences we used were not identical to Saffran et al., 1999, but had the same basic structure and level of complexity. One possible explanation for this negative result is that our sequences resemble melodic fragments known by the adults. However, this option seems unlikely because we

	SumSq	DF	MeanSq	F	P
Intercept	28.15	1.0	28.15	3086.21	0.00000
Between groups	0.02	1.0	0.02	2.32	0.13480
Error	0.39	43.0	0.01		
Within groups	0.01	1.0	0.01	0.86	0.35890
Interaction	0.00	1.0	0.00	0.18	0.67040
Error	0.72	43.0	0.02		

TABLE 5.7: Statistical analysis for Experiment 7. 2-ways ANOVA on the percentage of responses as dependent variable. The between factors is the familiarization phase condition (continuous/ pauses) and the within factor the force choice trial type (W vs. P/ W vs. S)

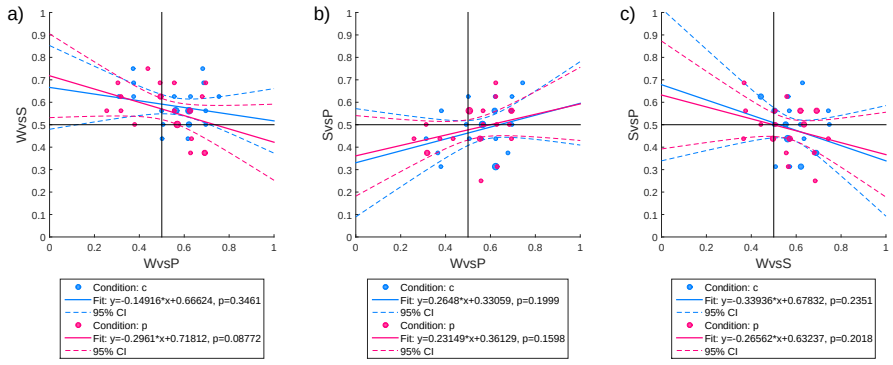


FIGURE 5.5: Individual answers for Experiment 7 for the three types of force choice trials plotted against each other. The size of each dot represents the number of subjects for that data point. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Answers for Words vs. Part-words force choice trials in the x axis, and for Words vs. Switch-words in the y axis. b) Answers Words vs. Part-words force choice trials in the x axis, and for Switch-words vs. Part-words in the y axis. c) Answers for Words vs. Switch-words force choice trials in the x axis, and Switch-words vs. Part-words in the y axis.

used two different familiarization streams (the words of one were created by displacing the tones that conformed the words of the other stream one position) and we did not find differences in performance between them.

A more plausible explanations is that the amount of familiarization was not enough. In our experiment participants heard the familiarization sequence for about 4 minutes, during which each tone sequence (word) was presented 80 times. In Saffran et al., 1999 the familiarizations lasted 7 minutes, during which each word was repeated 108 times; whereas in Abla, Katahira, and Okanoya, 2008 participants over-went 3 familiarization blocks each lasting 6.6 minutes during witch each word was repeated 40 times (120 repetitions in total). The duration of the tones was also slightly different between studies: 250 ms in ours, 330 ms in Saffran et al., 1999, and 550 ms in Abla, Katahira, and Okanoya, 2008; however, we

do not think this is the cause of the worst performance in our study. We did observe some degree of learning, which supports that the duration of the familiarization was indeed too short. A last element that may be responsible of the lack of learning, is that subjects were presented with three word classes during test blocks instead of two; yet, this was the same in Experiment 6, in which subjects performed significantly above chance.

If the cause of the failure in extracting the word like tone sequences was that the familiarization was too short, it implies that adults perform worst in segmenting non-linguistic than linguistic stimuli —the duration and structure of the familiarizations of Experiments 6 and 7 was the same. This difference in performance may be a consequence of adults more extensive experience with language. The experiment should be repeated with a longer familiarization in order to get conclusive results.

5.4 Adults segmenting linguistic visual stimuli. Experiment 8

In Experiment 6 we found evidence suggesting that once adults segments words from continuous speech they start encoding also the position of the syllables inside the words. We asked if this was a specific language mechanism or if it was a general mechanisms, and in Experiment 7 we tried to uncover what happens when participants have to extract the structure from auditory non linguistic stimuli. Here, instead, we ask what happens when the stimuli are linguistic, but are presented in the visual modality. To do so we replaced the syllables of Experiment 6 by written syllables that appear consecutively on the screen. One group of subjects was familiarized with a stream of syllables appearing on the screen one right after the other (continuous condition), whereas for a second group of subjects the screen remained empty for 50 ms between the syllables belonging to different words (pause condition).

If an abstract encoding is generated in the linguistic domain independently of the modality, we expect to find similar results than in Experiment 6. Instead if it differs between the auditory and visual domain we should observe that subjects prefer Words over Part-words and Switch-words (they succeed in the segmentation task), but they prefer Part-words over Switch-words.

5.4.1 Participants

Participants were the same than in Experiment 7 (see 5.3.1).

5.4.2 Stimuli

The familiarization streams and test items had identical distributional properties than in Experiment 6 (see 5.6.2), but the twelve syllables were replace by twelve

Stream	W	P	S
A	BASEZI	SEZIFE	DOSEZI
	DOKAVE	GADOKA	NIKAVE
	NITOGA	VENITO	FETOGA
	FEPIRO	PIROBA	BAPIRO
B	SEZIDO	DOKAVE	KAZIDO
	KAVENI	NITOGA	TOVENI
	TOGAFE	FEKAVE	PIGAFE
	PIROBA	ROBASE	SEROBA

TABLE 5.8: Stimuli for Experiment 8. Words, Part-words and Switch-words used in the experiment. Two streams and sets of words were created and participants were randomly assigned one. Syllables were presented written on the screen

written syllables (see Table 5.8). Each syllable was presented in white on a black background and staid on the screen for 500 ms. In the continuous condition syllables succeeded on the screen one after the other. In the Pause condition breaks of 50 ms were inserted between syllables of different words.

5.4.3 Procedure

The procedure was analogue to Experiments 6 (see 5.2.3).

During the training phase participants saw three syllables that succeed one after the other for 15 times, and afterwards they completed 9 test trials. During the test trials two three-items sequences were presented one after the other —one of them had the correct order— and participants had to choose the one they though was more likely part of the sequence they saw before. Participants repeated the training till having at least 75% of correct responses.

During the familiarization phase the stream of syllables was presented on the screen. In order to avoid perceivable starting and ending points, the luminance of the first 16 syllables was gradually increased and decreased for the last 16 syllables. The total duration of the familiarization was 8 minutes for the Continuous condition and 8 minutes 16 seconds for the Pause condition. During the test phase participants had to choose between two three-items sequences. As in previous experiments, the four words, part-words and switch-words were compared in three types of comparison: a) Word vs. Part-word b) Word vs. Switch-word c) Part–word vs. Switch word; in 16 trials per condition.

5.4.4 Apparatus and data acquisition

Idem than for Experiment 6. See 5.2.4.

5.4.5 Data Analysis

Idem than for Experiment 6. See 5.2.5.

Also in this case we did not find any effect or interaction for the familiarization stream (A/B) (P > 0.05), thus we did not include it as a factor in the final analysis reported here.

5.4.6 Results

The proportion of answers for each type of words are presented in Figure 5.6. Both, the t-test against chance level (see Table 5.9), and the 2-ways ANOVA (see Table 5.10) revealed that subjects extracted the words from the stream in both conditions: subjects preferred Words over Part-words (P < 0,05) and over Switch-words (P < 0,05) in both conditions. The only comparison did not survive after correcting for multiple comparisons was Words vs. Part-words in the continuous condition ($P_{un-corrected}$ = 0.0091). The 2-ways ANOVA revealed a main effect of familiarization phase (F(1,43) = 5.07, P < 0.05), and a main effect of type of test trial (F(1,43) = 7.51, P < 0.01), but no interaction (P > 0.05). Even if no interaction was present, we performed a multiple comparison post-hoc analysis using Tukey-Kramer correction, and we found a significantly higher preference for Words over Switch-words than over Part-words only in the continuous condition (P < 0.05) (see Table 5.11). In addition the pause and the continuous condition differed each other only for the Word vs. Part-words force choice trials. In Part-words vs. Switch-words force choice trials, participants did not show any preference in the pause condition (P > 0.05) and showed a marginally significant preference for Part-words in the continuous condition (P = 0.0583) (see Table 5.9).

As in the previous experiments we plotted the proportion of answers each subject gave for each type of force choice trial against each other (see Figure 5.7). When Word vs. Part-words force choice answers were plotted against Words vs. Switch-words answers, dots fell mainly in the upper right quadrant, indicating that subjects exacted the words. Subjects that performed at chance level in the Words vs. Part-words trials performed above chance in the Words vs. Switch-words trials, which is similar to the pattern we obtained in the experiment with linguistics auditory stimuli in the continuous condition. Diverging from the Experiment 6, in this experiment we did not find a correlation between Words vs. Part-words force choice trials (good segmentation) and Switch-words vs. Part-words trials (abstract encoding of position). Furthermore, Words vs. Switch-words and Switch-words vs. Part-words answers were negatively correlated, meaning that subjects with a high preference for Words over Switch-words, preferred Part-words over Switch-words.

	WvsP		**WvsS**		**PvsS**	
	c	**p**	**c**	**p**	**c**	**p**
P	0.0091	0.0002	0.0000	0.0000	0.0073	0.1700
P corr.	0.0730	0.0013	0.0000	0.0000	0.0583	1.3640

TABLE 5.9: Statistical analysis for Experiment 8. t-test against chance level for the percentage of responses for each type of force choice trial and familiarization condition. Bonferroni correction was applied to correct for multiple comparisons.

	SumSq	DF	MeanSq	F	P
Intercept	40.34	1.0	40.34	1291.55	0.00000
Between groups	0.16	1.0	0.16	5.07	0.02950
Error	1.34	43.0	0.03	1.00	0.50000
Within groups	0.16	1.0	0.16	7.51	0.00890
Interaction	0.01	1.0	0.01	0.64	0.42930
Error	0.91	43.0	0.02	1.00	0.50000

TABLE 5.10: Statistical analysis for Experiment 8. 2-ways ANOVA on the percentage of responses as dependent variable. The between factors is the familiarization phase condition (continuous/ pauses) and the within factor the force choice trial type (W vs. P/ W vs. S)

	Difference	StdErr	P	CI lower	CI upper
c: WvsP - WvsS	-0.10870	0.04297	0.01516	-0.19534	-0.02200
p: WvsP - WvsS	-0.05966	0.04393	0.18154	-0.14825	0.02890
WvsP: c - p	-0.10845	0.04680	0.02532	-0.20283	-0.01410
WvsS: c - p	-0.05941	0.04975	0.23898	-0.15975	0.04090

TABLE 5.11: Statistical analysis for Experiment 8. Multiple comparison post-hoc analysis. Tukey-Kramer correction was used.

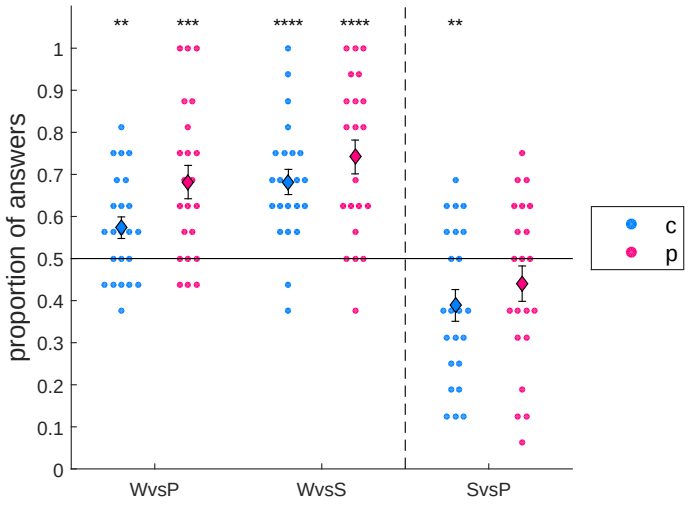


FIGURE 5.6: Answers for each type of force choice test trial after participants were familiarized with a stream of written syllables. Each dot represents the proportion of answers for one subject in a specific comparison. Results for the group familiarized with the continuous stream is plotted in blue, and with the stream with pauses in red.

5.4.7 Discussions

Adults succeed to extract the words, as demonstrated by their preference for Words over Part-words and Switch-words. Moreover, the pauses helped the segmentation. However, the pattern of results was different than in Experiment 6. Unlikely with speech, in this experiment we did not find an interaction between familiarization (pause / continuous) and force choice trial (W vs P / W vs S): as subjects get better in the segmentation they increase they preference for Words over both Part-words and Switch-words in equal way. Furthermore, participants did not prefer Switch-words over Part-words in any of the conditions.

These results entail that subjects keep relying on the co-occurrence of syllables to decide which sequence is more likely a word, even when they demonstrate a high performance in the segmentation task. Accordingly, when we looked to individual performance we did not find a positive correlation between preference for Words over Part-words and preference for Switch-words over Part-words, as we did in Experiment 6.

All together our results convey evidence that how information is encoded after segmentation is modality dependent. In both Experiment 6 and 8, the stimuli were linguistic, but only in the former experiment, in which speech was used, we observed results suggesting that the positional information of the syllables in the words becomes more relevant that the local links between syllables.

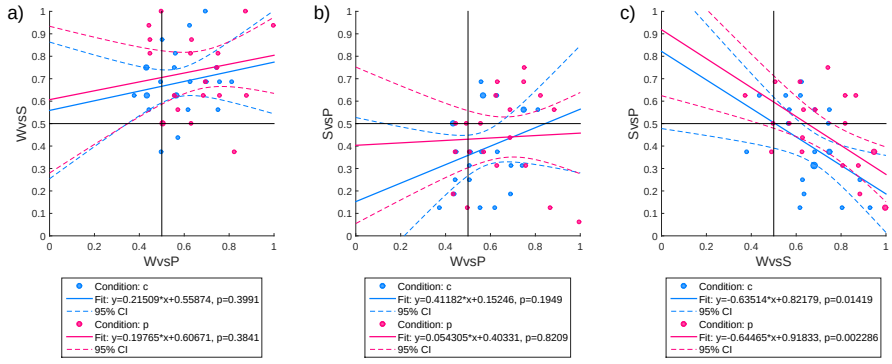


FIGURE 5.7: Individual answers for Experiment 8 for the three types of force choice trials plotted against each other. The size of each dot represents the number of subjects for that data point. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Answers for Words vs. Part-words force choice trials in the x axis, and for Words vs. Switch-words in the y axis. b) Answers Words vs. Part-words force choice trials in the x axis, and for Switch-words vs. Part-words in the y axis. c) Answers for Words vs. Switch-words force choice trials in the x axis, and Switch-words vs. Part-words in the y axis.

5.5 Adults segmenting non-linguistic visual stimuli. Experiment 9

In order to full-fill the picture we tested subjects in a new experiment using visual non-linguistic stimuli. The experiment was exactly as the previous one but written syllables were replace by arbitrary shapes. As before we tested two groups of subject, one in the continuous condition and one in the pause condition.

Based on Experiments 6 and 8 we expect participants to prefer Part-words over Switch-words, and in general terms similar results than in Experiment 8.

5.5.1 Participants

Participants were the same than in Experiment 6 (see 5.2.1).

5.5.2 Stimuli

For this experiment the written syllables were replace by twelve different shapes. All shapes were white and presented on a black background (see Figure 5.8).

5.5.3 Procedure

Idem than for Experiment 8. See 5.4.3.

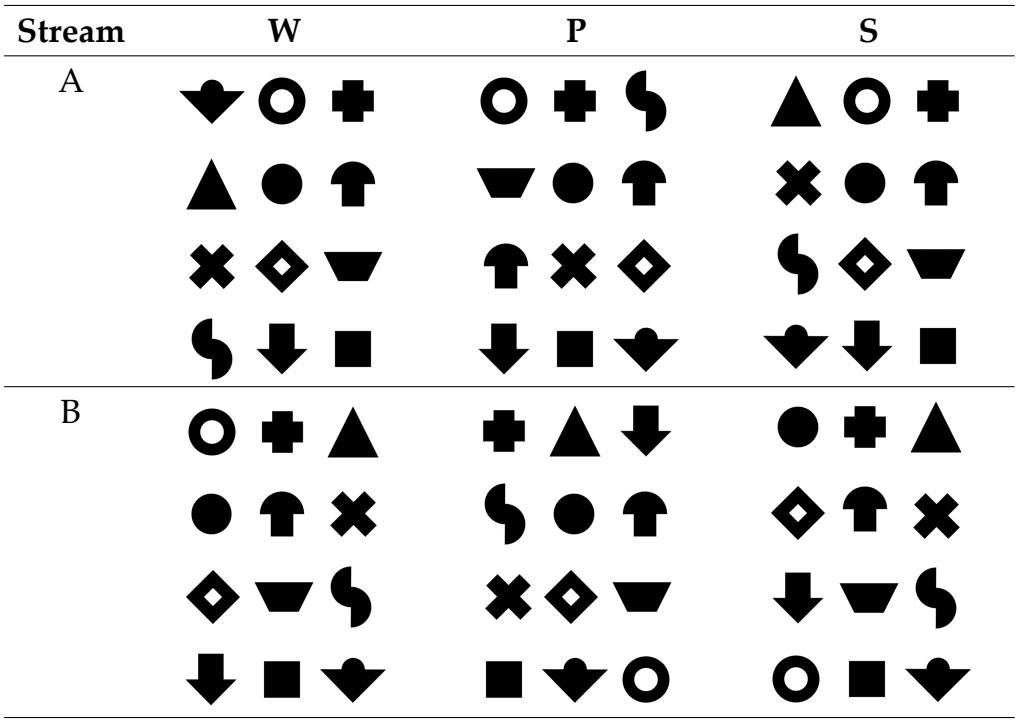


FIGURE 5.8: Stimuli for Experiment 9. Words (W), Part-words (P) and Switch-words (S) used in the experiment. Two streams and sets of words were created and participants were randomly assigned one.

5.5.4 Apparatus and data acquisition

Idem than for Experiment 6. See 5.2.4.

5.5.5 Data Analysis

Idem than for Experiment 6. See 5.2.5. Also in this case we did non find any effect or interaction for the familiarization stream (A/B) (P > 0.05), thus we did not include it as a factor in the final analysis reported here.

5.5.6 Results

The proportion of answers for each type of words are presented in Figure 5.9. As in the previous experiment, based on the t-test against chance level (see Table 5.9) and the 2-ways ANOVA (see Table 5.10), we can conclude that subjects segmented and extracted the sequences from the stream: they showed a preference for Words over both Part-words and Switch-words in both conditions. The only comparison that did not survive after correcting for multiple comparisons was Words vs. Part-words in the continuous condition ($P_{un-corrected}$ = 0.0323). The 2-ways ANOVA revealed a main effect of type of test trial (F(1,43) = 4.91, P <

0.05), but unlikely with written syllables we did not find an effect of familiarization phase (P > 0.05) and we did find an interaction (F(1,43) = 5.44), P < 0.05). By post-hoc multiple comparisons analysis using Turkey-Kramer correction, we uncovered that the origin of the interaction was a higher preference for Words over Switch-words than over Part-words in the continuous condition (P < 0.01), which was not present in the pauses condition (P > 0.05). The performance in the force choice trial Switch-words vs. Part-words was significantly below chance for both the continuous ($P_{un-corrected}$ = 0.0233) and pauses ($P_{un-corrected}$ = 0.0340) groups, but comparisons did not survive the multiple comparisons correction.

When we plotted the proportion of answers each subject gave for each type of force choice trial against each other (see Figure 5.10), we found a pattern of results very similar to the one using written syllables. In brief, subjects with a higher preference for Words over Part-words also showed a preference for Words over Switch-words, but when Part-words and Switch-words were confronted they preferred Part-words.

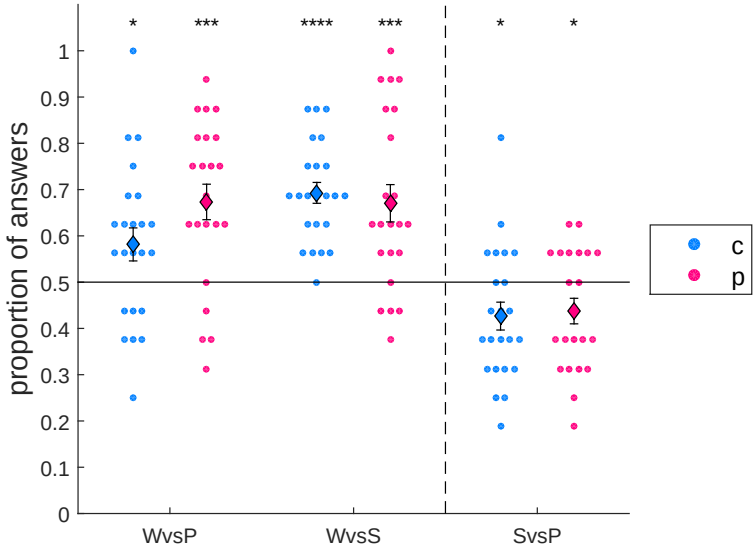


FIGURE 5.9: Answers for each type of force choice test trial after participants were familiarized with a stream of shapes. Each dot represents the proportion of answers for one subject in a specific comparison. Results for the group familiarized with the continuous stream is plotted in blue, and with the stream with pauses in red.

5.5.7 Discussions

The results of this experiment are very similar to the results of Experiment 8 as expected. The main difference is that the ANOVA did not reveal a significant effect of condition, but instead it showed a significant interaction. However, the

	WvsP		**WvsS**		**PvsS**	
	c	**p**	**c**	**p**	**c**	**p**
P	0.0323	0.0002	0.0000	0.0004	0.0233	0.0340
P corr.	0.2585	0.0015	0.0000	0.0030	0.1864	0.2730

TABLE 5.12: Statistical analysis for Experiment 9. t-test against chance level for the percentage of responses for each type of force choice trial and familiarization condition. Bonferroni correction was applied to correct for multiple comparisons.

	SumSq	DF	MeanSq	F	P
Intercept	38.54	1.0	38.54	945.96	0.00000
Between groups	0.03	1.0	0.03	0.66	0.42010
Error	1.75	43.0	0.04		
Within groups	0.07	1.0	0.07	4.91	0.03200
Interaction	0.07	1.0	0.07	5.44	0.02440
Error	0.58	43.0	0.01		

TABLE 5.13: Statistical analysis for Experiment 9. 2-ways ANOVA on the percentage of responses as dependent variable. The between factors is the familiarization phase condition (continuous/ pauses) and the within factor the force choice trial type (W vs. P/ W vs. S)

	Difference	StdErr	P	CI lower	CI upper
c: WvsP - WvsS	-0.11141	0.03424	0.00222	-0.18047	-0.04240
p: WvsP - WvsS	0.00284	0.03501	0.93571	-0.06777	0.07340
WvsP: c - p	-0.09177	0.05231	0.08650	-0.19727	0.01370
WvsS: c - p	0.02248	0.04567	0.62507	-0.06963	0.11460

TABLE 5.14: Statistical analysis for Experiment 9. Multiple comparison post-hoc analysis. Tukey-Kramer correction was used.

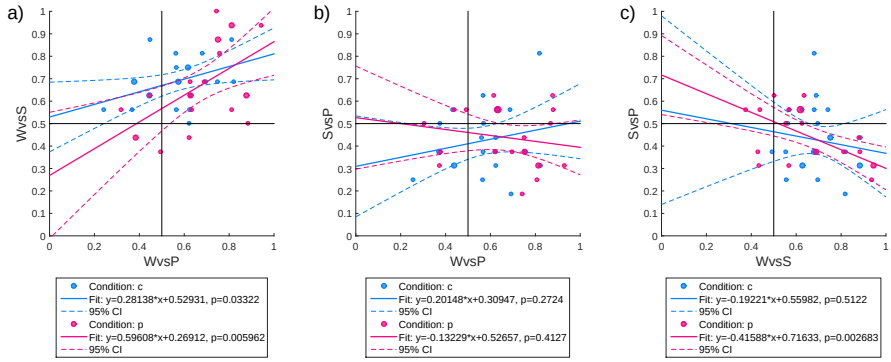


FIGURE 5.10: Individual answers for Experiment 9 for the three types of force choice trials plotted against each other. The size of each dot represents the number of subjects for that data point. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Answers for Words vs. Part-words force choice trials in the x axis, and for Words vs. Switch-words in the y axis. b) Answers Words vs. Part-words force choice trials in the x axis, and for Switch-words vs. Part-words in the y axis. c) Answers for Words vs. Switch-words force choice trials in the x axis, and Switch-words vs. Part-words in the y axis.

origin of the interaction is different than in Experiment 6. In fact, subjects under both, the continuous and pause condition, prefered Part-words over Switch-words. Further more, when we looked to individual data by plotting the Words vs. Switch-words and Switch-words vs Part-words force choice trials, points appeared mainly in the lower right quadrant, indicating that, as with written syllables, subjects that recognized the Word sequences still preferred Part-words over Switch-words.

Results of this last experiment confirmed that for visual stimuli sequentially presented, which it is retained is the association between pairs of elements. Even when the participants extracted the word like sequences we did not obtain results compatible with a positional encoding of items inside shorter sequences (words). In other words, we did not find evidence of a more abstract —less local— representation of the sequences conforming the stream as we did for speech.

5.6 5-month-old segmenting linguistic auditory stimuli. Experiment 10

In the previous experiments we tested our initial hypothesis in adults across domains. Notwithstanding, these experiments do not provide a full picture. We cannot say from the adults results if this peculiarity observed in speech processing is innate or a consequence of the long experience with language that adults have. In language position is crucial: words can be exchanged in a phrase if they

belong to the same syntactic category, and morphology appears always in specific positions. For example articles and pronouns in Italian tend to appear at the beginning of intonational phrases, and verbs morphology is done at the end of the verbs. These common patterns in language could render adults more likely to judge a sequence of syllables that conserves positional information as part of the language they were familiarized during the experiment. To investigate this, we decided to test the same in infants.

We aimed to test infants with as little as possible language experience, and we needed a measure capable of reflecting a modulation by three levels: Words, Part-words and Switch-words. Segmentation in infants has been traditionally tested using the head turning preference procedure (Saffran, Aslin, and Newport, 1996), but this protocol enables to compare only two conditions (differences in looking time between more conditions are not usually obtained). In our experiment we implemented a new methodology based in the classical head turning preference procedure, but adapted to eye tracking, which enables to record pupil dilation. Our methodology provides two main advantages. First, it does no requires coding the data, which reduces the amount of work and makes results less subject to biases from the experimenter. Second, both the looking time and the pupil diameter can be recorded. Pupil dilation is an automatic response that mainly depends of the level of luminance, but that also correlates with arousal, attention and cognitive load in adults (e.g. Hess and Polt, 1960; Hess and Polt, 1964; Kahneman and Beatty, 1966, for a review see Laeng, Sirois, and Gredeback, 2012). The automatic nature of the response makes it a very suitable dependent measure for the study of infant cognition, but that started to take relevance in the field only recently (Jackson and Sirois, 2009; Gredebäck and Melinder, 2010; Sirois and Jackson, 2012; Lewkowicz and Hansen-Tift, 2012). Even if pupillometry requires a rigorous control for the level of luminance, if well designed paradigms are achieved, it provides another independent measure reflecting surprise or attention, to the classically used looking times. Furthermore, pupil size is measured along time, thus a time course of the measure with changes locked to the events is obtained.

Pupillometry appears as a suitable measure for our goal, which requires a measure that reflects differences in attention (or cognitive load) to the different types of words after the familiarization. The automatic nature of pupil dilation may reflect a modulation for the three words categories. Additionally, our auditory task does not require the presentation of visual stimuli, hence controlling luminosity does not represent an issue.

We tested 5-month-old, the younger age we could test using eye tracker. Infants were tested using the exact same stimuli we used for adults in Experiment 6. One group of infants was familiarized with the continuous stream and another group of infants with the stream of pauses. After the familiarization phase, infants underwent 12 test trials (four per condition), in which they listen either a Word, a Part-word or a Switch-word, and the looking time and pupil diameter

towards an attractor was measured.

If infants process and encode the speech as adults, we expect to observe: 1. shorter looking times and smaller pupil size for Words than for Part-words and Switch-words, as an indication that they segmented the stream; 2. longer looking times and bigger pupil dilation for the Part-words than for Switch-words, denoting Switch-words are more familiar; 3. a positive correlation between infant level of surprise for Part-words respect to Words (how well they segmented the words), and the differential surprise between Part-words and Switch-words (how much they consider Switch-words closer to Words than to Part-words); 4. stronger effects for the group of infants familiarized with the stream with pauses, if pauses help infants in the segmentation.

5.6.1 Participants

All participants were healthy full-term infants. The continuous condition included 23 infants (10 females; mean age 5.1 month; range 4.3-5.6 month; SD 0.3 month). The pause condition included other 23 infants (8 males; mean age 5.1 month; range 4.5-5.7 month; SD 0.3 month). Additional infants came to the lab to participate to the experiments, but did not completed the task due to fussiness (Continuous condition n = 10, Pause condition n = 8); whereas other infants achieved to do the task but were excluded from the final analysis due to poor data quality —not enough good trials (Continuous condition n = 3, Pause condition n = 2). Parents of the infants firmed a consent form, and the Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

5.6.2 Stimuli

The auditory stimuli were identical to the auditory stimuli of Experiment 6. Infants were randomly assigned to one or the other group.

The visual stimuli were an attractor and a central fixation cross, presented on a white background. In order to maintain the iso-luminance the fixation cross was created by scrambling the attractor. The attractor was spinning, whereas the fixation cross rotated back and forward rapidly. The size of both stimuli was 200x200 pixels. The attractor appeared either on one or the other side of the screen at 256 pixels from the the border, and the fixation cross in the center of the screen (see Figure 5.11).

5.6.3 Procedure

Before the experiment infants gaze was calibrated using five fixation points (at the corners and center of the screen) with a calibration protocol included in the eye tracker.

The experiment started with a familiarization phase in which the familiarization stream was played (see Figure 5.11). The stream was ramped up and down

during the first and last 6 s. To keep the attention of the infants and to familiarize them also with the visual stimulus, the attractor was presented alternating from one to the other side of the screen at intervals of varying length between 6 and 12 seconds. The total duration of the familiarization was 4 minutes.

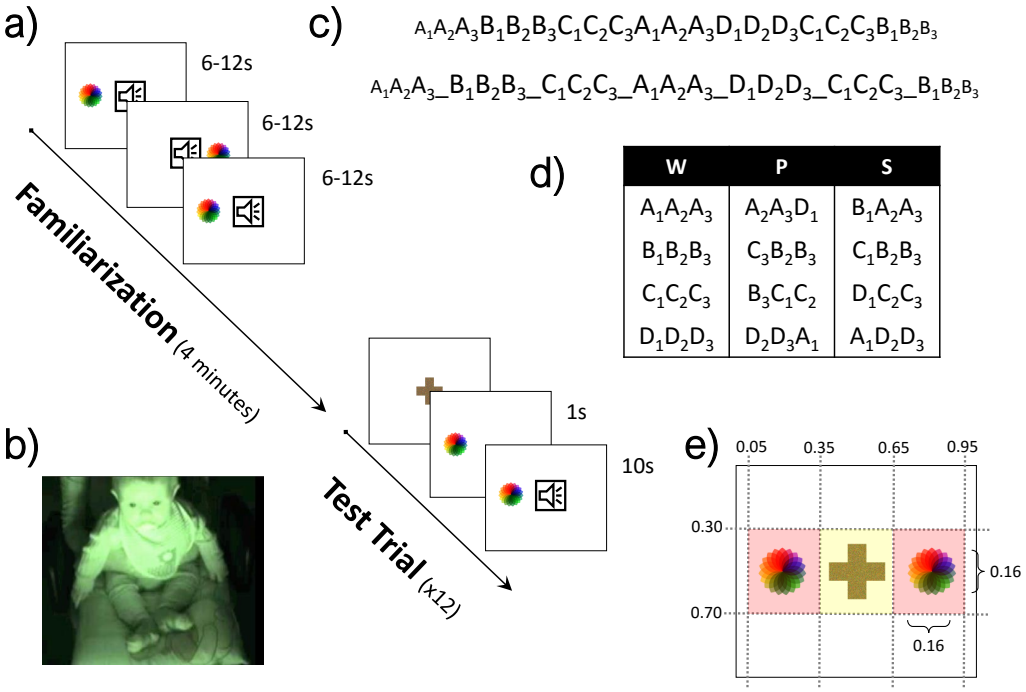


FIGURE 5.11: a) Time line of the experiment. b) Example of an infant seated on her parent's lap during the experiment. c) Familiarization streams were created by randomly concatenating 4 three-syllabic words. In the stream with pauses, 25 ms pauses were inserted between words. d) Structure of the Words, Part-words and Switch-words used during test trials. e) Regions of interest on the screen .

Twelve test trials —four per condition, were presented after the familiarization. Test trials start with a fixation cross. As soon as the infants fixate the cross, the attractor appears randomly at one or the other side of the screen, and after one second the sound starts. The test item is repeated during a total duration of 10 seconds, separated by silences of varying length between 500 and 800 ms.

5.6.4 Apparatus and data acquisition

Infants were tested in a dark booth. Only one of the parents stayed inside the booth during the experiment. The infant sat at her parent's lap, at approximate 60 cm from the screen, and parents were asked to wear opaque glasses.

Infants gaze was recorded using a TOBII T120 eye-tracker at 60Hz. The eye-tracker was integrated to a 15" monitor, resolution 1280x1024, where the stimuli were presented using a iMAC 10,1 computer. Auditory stimuli were presented via

two loudspeakers located symmetrically behind the screen. The audio was presented in stereo mode. The experiment was written in MATLAB, using the Psychophysics Toolbox extensions (Brainard, 1997). Infants behaviour was recorded with a camera hidden behind the screen.

5.6.5 Data Analysis

We calculated the fixations to the attractor and the pupil dilation, the two dependent measures that we expected to be modulated by the sound.

Gaze data. To estimated the fixations we followed a pre-processing procedure based on Wass, Smith, and Johnson, 2013. We compared the duration of the first fixation to the attractor with an ANOVA. The steps of the analysis are described in detail below. See Figure 5.12 for an example of a trial.

1. *Filtering*. To smooth the signal we applied a bilateral filter algorithm (Durand and Dorsey, 2002). This algorithm removes jitter but preserves saccades. We used, $\sigma_r = 0.1$ for the range parameter, and $\sigma_t = 4$ for the time parameter.

2. *Interpolation.* When data from only one eye was available, we used it to estimate the coordinates for the other eye, by linearly interpolating the time course of the distance between eye. Afterwards, we interpolated missing data samples for gaps shorter than 9 samples (150 ms). The value was chosen considering that the time to plan a saccade is 100-130 ms, making unlikely the sequence saccade-fixation-saccade in an interval of 150 ms (Saez de Urabain, Johnson, and Smith, 2015; Wass, Smith, and Johnson, 2013). Data were interpolated by continuing forward and backward the average value of the previous 9 and the subsequent 9 good data point around the gap. The coordinates of both eye were interpolated independently.

3. *Gaze calculation.* We calculated the gaze as the middle point between the position of the left and right eye.

4. *Fixations.* We identified a fixation as a period in time of at least 150 ms delimited by a gaze velocity bigger than $50°/s$. We identified a fixations towards a specific object when the gaze coordinates were inside a defined region of interested. We defined three regions of interest: the fixation cross (x_1 = 0.35, x_2 = 0.65, y_1 = 0.30, y_2 = 0.70); the attractor on the left (x_1 = 0.05, x_2 = 0.35, y_1 = 0.30, y_2 = 0.70; all in relative units to the screen size), and the attractor on the right (x_1 = 0.65, x_2 = 0.95, y_1 = 0.30, y_2 = 0.70; all in relative units to the screen size). All the positions are expressed in relative units (see Figure 5.11).

5. *Inclusion criteria.* We only included trials in which: a) the infant fixated the fixation cross before the attractor appeared b) the infant fixated the attractor in the interval [-0.2s, 0s] before the sound started c) the infant looked at least 1.5 seconds to the attractor after the sound started (we considered the looking to the attractor over when the infant looked away for more than 1 second). d) the infants provided at least one good trial per condition. In the continuous conditions infants provided an average of 3.34 trials (SD = 0.88) in the words condition, 3.22 (SD = 0.95) in the part-words condition, and 3.52 (SD = 0.73) in the switch-words condition. In the pause conditions infants provided an average of 3.39 trials (SD = 0.84) in the words condition, 3.61 (SD = 0.72) in the part-words condition, and 3.65 (SD = 0.52) in the switch-words condition. The third criteria was set for two main reasons. On one hand to avoid including trials in which infants were not paying attention. On the other hand, infants have to look to the attractor for a minimum amount of time to make possible to use the pupil dilation as a dependent measure.

6. *Statistical analysis.* We used the duration of the first fixation as dependent measure. We run an 2 way-ANOVA with factor Condition (continuous/pause) as between subjects factor and Type of test trial (W/P/S) as within subjects factor.

Pupil data. We obtained the pupil diameter time course from the appearance of the attractor, and we set time zero at the onset of the sound. The pupil dilation towards the sound was used as dependent measure to performed the statistical analysis. A detailed explanation of the analysis is presented below.

1. *Low pass filter.* To reduce variability and flikering we low pass filtered the data at 4.8 Hz —ratio between the sampling frequency and the cut of frequency 12.5, (Jackson and Sirois, 2009).

2. *Interpolation.* The left and right pupil sizes are highly correlated, hence when pupil data of only one eye was available we used it to estimate the pupil size of the other eye, by linearly interpolating the time course of the difference in pupil diameter between eye. Subsequently, we interpolated missing data samples, usually due to blinking. We did a linear interpolation from the previous 3 and subsequent 3 good data points around the gap. If missed data points were at the begging/end of the trial we extrapolated the first/last good data point till the begging/end of the trial.

3. *Pupil dilation to the attractor.* We estimated the pupil diameter as the mean pupil diameter between eye. For each trial data were linearly detrended. Afterwards we calculated the percentage change respect ot a baseline. As

baseline we used the mean value in the period [-0.2s, 0s], while they were looking to the attractor.

4. *Statistical analysis.* We used the average pupil diameter change in the period [0s, 1.5s], as dependent measure. We run an 2 way-ANOVA with factor Condition (continuous/pause) as between subjects factor and Type of test trial (W/P/S) as within subjects factor.

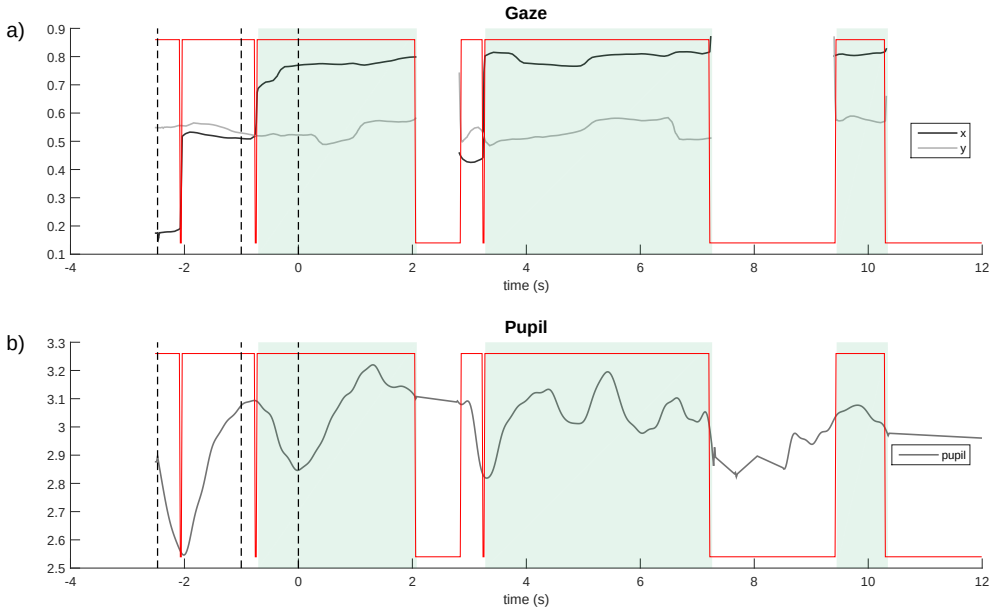


FIGURE 5.12: Example of a trial for one subject during Experiment 10. Dotted lines indicated the appearance of the fixation cross, of the attractor and the beginning of the sound. The red line marks the periods where fixations were detected. The green shaded area are the moments in which the gaze was in the region of interest of the attractor. a) Gaze coordinates b) Pupil diameter.

Individual data inspection. We wanted to explore if as adults, infants preference for Part-words or Switch-words depends of their individual performance in the segmentation task. In order to explore the data we plotted the differences in pupil dilation between conditions against each other. Specifically we computed: (i) Part-words - Words vs. Switch-words - Words; (ii) Part-words - Words vs. Part–words - Switch-words; (iii) Switch-words - Words vs. Part-words - Switch-words.

5.6.6 Results

The time course of the pupil dilation for the three types of test trials is shown in Figure 5.13.

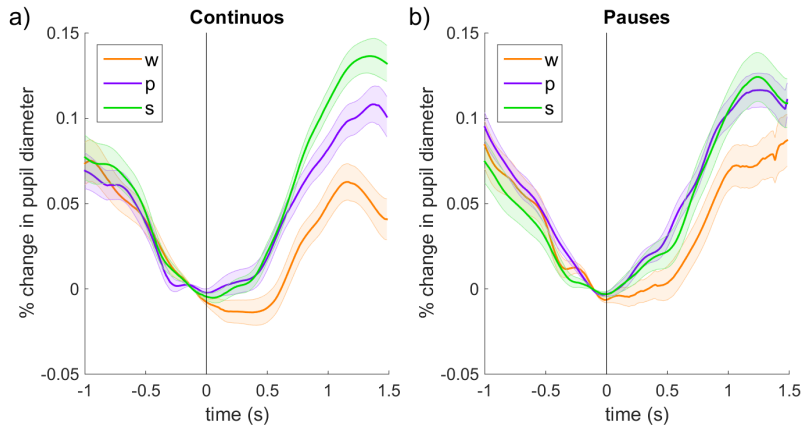


FIGURE 5.13: Pupil dilation time course for Experiment 10 for the three test trial types: Words (orange), Part-words (violet) and Switch-words (green). The shaded area represents standard error. Time zero is the beginning of the sound.

We performed the statistical analysis on the duration of the first fixation and the mean pupil dilation. Results are presented in Figure 5.14. By adding the familiarization stream (A/B) as a between subject factor in the analysis we did not find any effect of it or interaction for any of the measures ($P > 0.05$), therefore we did not consider it in the final analysis. The results of the 2-ways ANOVA for the two dependent measures are shown in Table 5.15. The duration of the first fixation revealed a significant effect of test trial type ($F(2,88) = 6.29$, $P < 0.005$), a marginally significant effect of familiarization phase ($F(1,44) = 3.63$, $P = 0.063$), and not interaction. Post hoc multiple comparison Tukey-Kramer corrected test, revealed that the first look to the attractor when Switch-words were presented was longer than when either the Words ($P < 0.01$) or Part-words ($P < 0.05$) were presented. On the mean pupil dilation we found a main effect of test trial ($F(2,88) = 9.39$,$P < 0.0005$), and no effect of familiarization phase or interaction ($P > 0.05$). Post hoc multiple comparison Tukey-Kramer corrected test, revealed that the pupil dilation was significantly smaller for Words than for Part-words ($P < 0.005$), and Switch-words ($P < 0.005$) (See Table 5.16).

By inspecting the individual data (see Figure 5.15) we could observe that subjects that showed a bigger pupil dilation for Part-words than for Words also showed a bigger pupil dilation for Switch-words than for Words, which is consistent with the fact that infants segmented the stream. However, we did not find consistent evidence that infants performing better show less surprise for Part-words than for Switch-words. Pupil dilation for Part-words and Switch-words were very similar for all subjects, and when its difference was plotted against Part-words - Words we did not observe that good performers had a bigger pupil dilation for Part-words than for Switch-words (see Figure 5.15 b)).

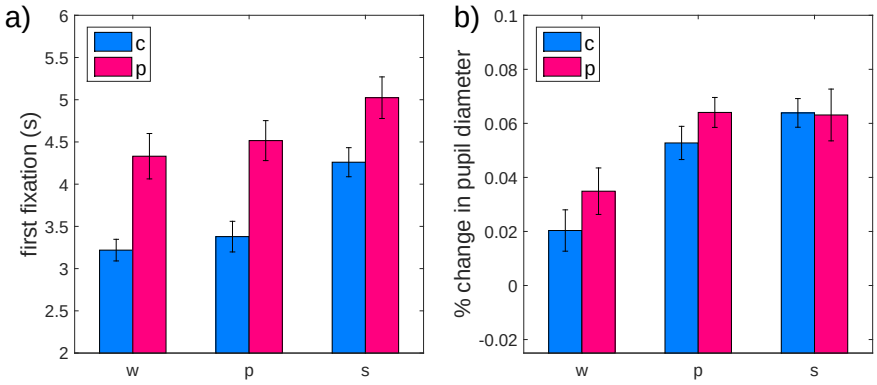


FIGURE 5.14: Duration of the first fixation (a) and pupil dilation (b) for Experiment 10. Error bars represent within subjects standard errors.

		SumSq	DF	MeanSq	F	P
First fixation	Intercept	2344.16	1	2344.16	244.51	0.00000
	Condition	34.80	1	34.80	3.63	0.06330
	Error	421.83	44	9.59		
	Test type	19.38	2	9.69	6.29	0.00280
	Interaction	1.00	2	0.50	0.33	0.72330
	Error	135.51	88	1.54		
Pupil dilation	Intercept	0.34279	1	0.34279	15.77	0.00030
	Condition	0.00241	1	0.00241	0.11	0.74060
	Error	0.95612	44	0.02173		
	Test type	0.03464	2	0.01732	9.39	0.00020
	Interaction	0.00150	2	0.00075	0.41	0.66800
	Error	0.16235	88	0.00184		

TABLE 5.15: Statistical analysis for Experiment 10. 2-ways ANOVA on the duration of the first fixation to the attractor and pupil dilation. The familiarization stream (continuous/ pauses) is a between subjects factor, whereas the test trial type is a within subjects factor (w/p/s)

		Difference	StdErr	P	CI lower	CI upper
First fixation	p-s	-0.6947	0.2597	0.02766	-1.3247	-0.0650
	p-w	0.1722	0.2572	0.78226	-0.4516	0.7960
	s-w	0.8670	0.2593	0.00474	0.2380	1.4960
Pupil dilation	p-s	-0.0051	0.0078	0.79248	-0.0240	0.0140
	p-w	0.0308	0.0084	0.00188	0.0104	0.0510
	s-w	0.0359	0.0104	0.00367	0.0105	0.0610

TABLE 5.16: Statistical analysis for Experiment 10. Multiple comparison on the duration of the first fixation and pupil dilation. Tukey-Kramer correction was used

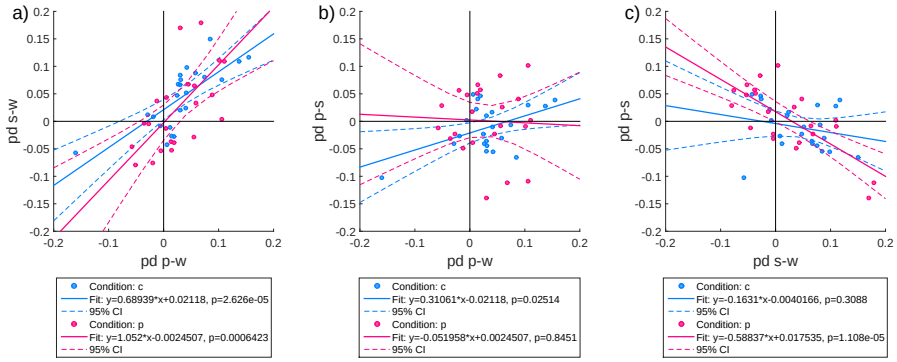


FIGURE 5.15: Individual differences for Experiment 10 for the three comparisons plotted against each other. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Part-words - Words in the x axis, and Switch-words - Words in the y axis. b) Part-words - Words in the x axis, and Part-words - Switch-words in the y axis. c) Switch-words - Words in the x axis, and Part-words - Switch-words in the y axis.

5.6.7 Discussions

In the current experiment we asked two main questions. First, what is more important in the representation 5-month-old create of words extracted from continuous speech: the transition probabilities between syllables or the positional information of the syllables in the word. Second, if subtle pauses marking the words enhance the segmentation and modify how infants encode the words.

Before being able to answer these questions we needed evidence that infants extracted the words. Experiment 10 revealed a modulation of the duration of the first fixation and of the pupil dilation by the type of word presented. The first fixation was longer for Switch-words than for Part-words and Words, whereas the pupil dilation was bigger toward Part-words and Switch-words than for Words. Both measures show that infants were more familiar or have less interest in the Words, but they differed regarding Part-words and Switch-words: the first fixation was longer for Switch-words than for the two other word classes; whereas the pupil dilation was bigger for both Part-words and Switch-words than for Words but no difference was observed between them. The origin of this discrepancy may be related to the nature of the two measures. Pupil dilation is an automatic response, thus easily modulated in a more continuous way; whereas looking times probably reflect a higher order cognitive process, which may be the reason why differences in looking times between more than two conditions are hard to get. We speculate that infants look longer towards the category that recalls more their attention between all the presented categories. If this is the case, our result indicate that infants have more interest in the Switch-word —items that they have never heard but which syllables appeared in the same position that actual words; than

in Part-words —sequences that were in the stream but do not have the structure of words. It is important to notice that equal looking times toward Part-words and Words cannot be attributed to infants not computing the TPs. The pupil data clearly shows that they are able to do so, otherwise we would not have observed a bigger pupil dilation for both Part-words and Switch-words, than for Words.

Summarizing, the pupil data confirms that infants segmented the stream, and the longer looking times toward Switch-words than toward Part-words, together with the absence of a difference in the pupil dilation between these two categories, indicates that infants found Switch-words more interesting than Part-words. The results clearly show that 5-month-old were able to segment the stream.

One of the aims of the study was to see if infants had a preference for Part-words or for Switch-words. The longer looking times toward Switch-words are not easily interpretable, because we do not know if infants are showing a novelty or a familiarity effect. In fact, it is well documented that when infants are still learning from a stimulus show a familiarity effect that later switches to novelty (Hunter, Michael A; Ames, 1988). The pupil dilation in this aspect is easier to interpret but unfortunately it was not different between Part-words and Switch-words.

Despite we did not find significant differences for the pupil dilation between Part-words and Words for the group as a whole, it could still be the case that the better performers (infants showing a bigger pupil dilation for Part-words than for Words) present bigger pupil dilation for Part-words than for Switch-words; whereas bad performers, who are still computing the distribution of the syllables and do not have a representation of the words, show a bigger pupil dilation for Switch-words than for Part-words. However, our data did not reveal consistent evidence to support this hypothesis neither, which makes tempting to conclude that infant behaviour diverges from adult behaviour in Experiment 6, who preferred Switch-words over Part-words. However, because we did not find a modulation of the pupil dilation between Part-words and Switch-words this conclusion is not strictly proper. If infants generate a representation of the Words only based on transitional probabilities, we should observe bigger surprise for Switch-words than Part-words, and this was not the case. A new experiment should be conducted with a longer familiarization or a simpler structure in order to see if differences between Part-words and Switch-words emerge.

Regarding the role of the pauses, we did not find a significant effect or interaction for the pupil dilation. The marginally significant effect we obtained for the first fixation, is likely a consequence of individual difference between infants in the two groups. There is not reason to think that a familiarization with pauses signalling the edges should induce longer looking times across all the conditions. Seemingly, the pauses do not help 5-month-old infants in the segmentation task as they do for adults. A first possible explanation is that because of the smaller

experience infants have with language, they are less sensitive to these disruptions; but considering that numerous studies (Hirsh-Pasek et al., 1987; Nelson et al., 1989; Gerken, Jusczyk, and Mandel, 1994; Nazzi, Jusczyk, and Johnson, 2000; Soderstrom et al., 2003; Christophe et al., 1994, and Experiments 2 5 in this thesis) show that even very young infants are particularly sensitive to prosodic cues, we consider this option unlikely. A second explanation is that infants' limitation for the task is not the segmentation by it self —regardless if based on the computation of the distributional cues, or on the perception of prosodic boundaries— but there are other limiting factors operating. For example one possible limiting factor could be memory. If this is the case, providing additional cues for the segmentation will not turn in a better performance. In order to see an effect of the pauses the number of syllables as well as the duration of the familiarization phase should be reduced.

5.7 5-month-old segmenting non linguistic auditory stimuli. Experiment 11

In Experiment 10 we tested how 5-month old segment and encode speech using pupils dilation as dependent measure. In order to see if any aspect of the processing and encoding was language specific we tested another two groups of infants using the same protocol, but with non linguistic stimuli. We used the exact same stimuli than we used for adults in Experiment 7.

Given that we did not find evidence of a positional encoding by infnats for speech, we do not expect to find it in the current experiment. Despite in Experiment 7 adults failed to extract the sequences of tones, previous experiments report success in this task in both infants (Saffran et al., 1999; Kudo et al., 2011) and adults (Saffran et al., 1999; Abla and Okanoya, 2009). Hence it may be the case that 5-month-olds succeed anyhow, maybe as consequence of less specialization on speech processing.

5.7.1 Participants

All participants were healthy full-term infants. The continuous condition included 23 infants (11 females; mean age 5.3 month; range 5.0-5.7 month; SD 0.2 month). The pause condition included other 23 infants (6 females; mean age 5.3 month; range 4.1-5.8 month; SD 0.4 month). Additional infants come to the lab to participate to the experiments, but did not complete the task due to fussiness (continuous condition n = 10, pause condition n = 12); whereas other infants achieved to do the task but were excluded from the final analysis due to poor data quality

—not enough good trials (continuous condition n = 4, pause condition n = 3). Parents of the infants firmed a consent form, and the Ethical Committee of the Scuola Internazinale Superiore di Studi Avanzati approved the study.

5.7.2 Stimuli

The auditory stimuli were identical to the auditory stimuli of Experiment 7 (see 5.3.2).

The visual stimuli were identical to Experiment 10.

5.7.3 Procedure

Idem than for Experiment 10. See 5.6.3.

5.7.4 Apparatus and data acquisition

Idem than for Experiment 10. See 5.6.4.

5.7.5 Data Analysis

Idem than for Experiment 10. See 5.6.5.

In the continuous conditions infants provided an average of 2.87 trials (SD = 1.39) in the words condition, 2.52 (SD = 1.08) in the part-words condition, and 3.17 (SD = 1.07) in the switch-words condition. In the pause conditions infants provided an average of 3.30 trials (SD = 1.11) in the words condition, 3.17 (SD = 1.02) in the part-words condition, and 3.43 (SD = 0.90) in the switch-words condition.

5.7.6 Results

The time course of the pupil dilation for the three types of test trials is shown in Figure 5.16.

As in the previous experiment, we run the statistical analysis on the duration of the first fixation and the mean pupil dilation. Results are presented in Figure 5.17. Neither in this experiment there was an effect of the familiarization stream (A/B) for any of the measures ($P > 0.05$), thus we did not consider it as a factor. The results of the 2-ways ANOVA for the two dependent measures are shown in Table 5.17. We did not find significant results for the duration of the first fixation ($P > 0.05$); yet, the mean pupil dilation revealed a main effect of test trial ($F(2,88) = 6.11$, $P < 0.005$), a marginally significant effect of familiarization condition ($F(1,44) = 3.14$, $P = 0.0832$), and no interaction ($P > 0.05$). Post-hoc multiple comparison Tukey-Kramer corrected test, revealed that the pupil dilation for Switch-words was bigger that for Words ($P < 0.05$) and Part-words ($P < 0.01$) (See Table 5.18).

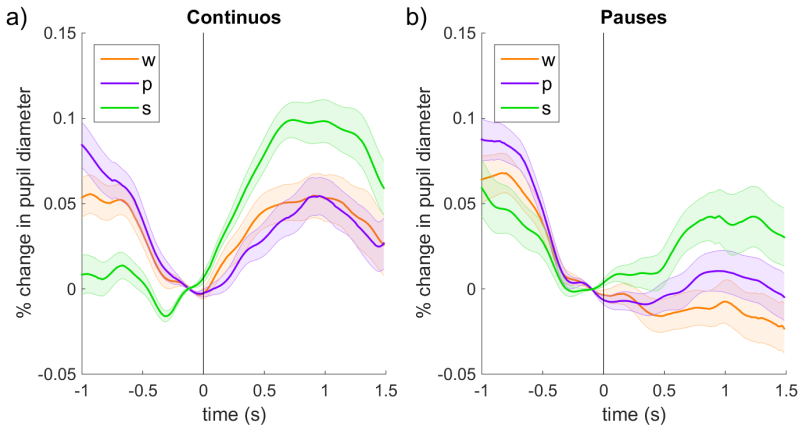


FIGURE 5.16: Pupil dilation time course for Experiment 11 for the three test trial types: Words (orange), Part-words (violet) and Switch-words (green). The shaded area represents standard error. Time zero is the beginning of the sound.

As in the previous experiment we did not find positive evidence of bigger surprise for Part-words than Switch-words in infants performing better in the task. (see Figure 5.18).

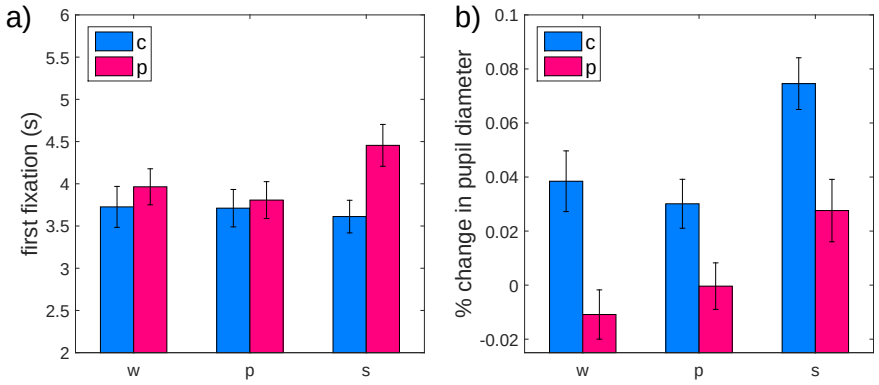


FIGURE 5.17: Cumulative looking time (a), duration of the first fixation (b) and pupil dilation (b) for Experiment 11. Error bars represent within subjects standard errors.

5.7.7 Discussions

In the current experiment we tested 5-month-old segmentation capacities of a continuous or separated by subtle pauses flow of tones. Above all we were interested on which aspects of the segmented word like sequences infants encode, hence we needed evidence that infants were actually able to segment the stream. Results from the pupil data revealed a significantly higher dilation toward Switch-words than toward Part-words and Words, but we did not find a significant difference between these two last conditions. Meanwhile, looking times did not reveal

		SumSq	DF	MeanSq	F	P
First fixation	Intercept	2076.90	1	2076.90	221.06	0.00000
	Condition	5.31	1	5.31	0.57	0.45610
	Error	413.39	44	9.40		
	Test type	1.81	2	0.90	0.52	0.59350
	Interaction	3.63	2	1.81	1.05	0.35300
	Error	151.45	88	1.72		
Pupil dilation	Intercept	0.09747	1	0.09747	4.97	0.03090
	Condition	0.06162	1	0.06162	3.14	0.08320
	Error	0.86261	44	0.01960		
	Test type	0.04145	2	0.02073	6.11	0.00330
	Interaction	0.00243	2	0.00121	0.36	0.70050
	Error	0.29860	88	0.00339		

TABLE 5.17: Statistical analysis for Experiment 11. 2-ways ANOVA on the cumulative looking to the attractor, duration of the first fixation and pupil dilation. The between groups factor is the familiarization stream (continuous/ pauses), whereas the within subjects factor is the test trial type (w/p/s)

		Difference	StdErr	P	CI lower	CI upper
First fixation	p-s	-0.2741	0.2676	0.56584	-0.9232	0.3750
	p-w	-0.0862	0.2750	0.94733	-0.7534	0.5810
	s-w	0.1878	0.2778	0.77858	-0.4861	0.8620
Pupil dilation	p-s	-0.0362	0.0118	0.00975	-0.0647	-0.0080
	p-w	0.0011	0.0112	0.99490	-0.0262	0.0280
	s-w	0.0373	0.0133	0.02049	0.0049	0.0700

TABLE 5.18: Statistical analysis for Experiment 11. Multiple comparison on the cumulative looking to the attractor, duration of the first fixation and pupil dilation. Tukey-Kramer correction was used.

any difference between any of the word categories. The higher pupil dilation to Switch-words indicates that infants were tracking the co-occurrence of the sounds and they could distinguish between a sequence of tones they had never heard —TPs equal to zero— and a sequence they had heard. Nevertheless, they could not distinguish between Words and Part-words —two sequences with all TPs bigger than zero. Probably, as for adults in Experiment 7 the familiarization was not long enough, and presumably we did not observe an effect in the looking time because this behavioural response requires a better performance in segmentation.

This result, together with Experiment 7, contradicts previous studies showing that both adults (Saffran et al., 1999; Abla and Okanoya, 2009) and infants (Saffran et al., 1999; Kudo et al., 2011) are able to segment a continuous stream of tones. Saffran et al. (Saffran et al., 1999) did not find a difference in performance between the two type of stimuli as we did, which may be consequence of the longer familiarization they used. If the segmentation performance in their experiment was at the selling level, by reducing the duration of the familiarization we may had evidenced better segmentation abilities, in both adults and 5-month-old, for

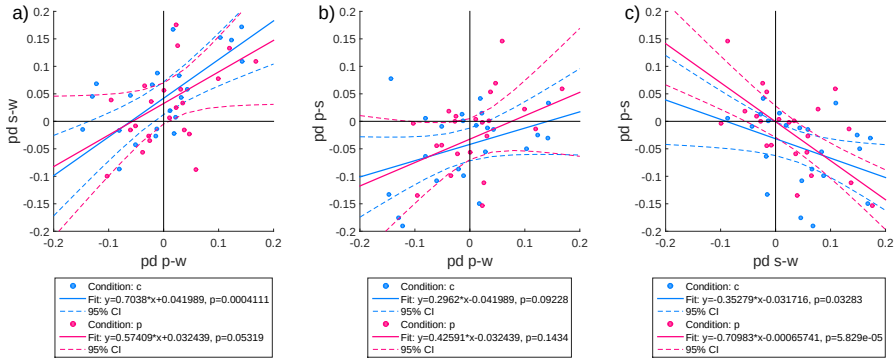


FIGURE 5.18: Individual differences for Experiment 11 for the three comparisons plotted against each other. The continuous condition is plotted in blue and the pause condition in red. Linear models fittings and confidence interval are reported. a) Part-words - Words in the x axis, and Switch-words - Words in the y axis. b) Part-words - Words in the x axis, and Part-words - Switch-words in the y axis. c) Switch-words - Words in the x axis, and Part-words - Switch-words in the y axis.

speech than for non linguistic auditory stimuli like tones.

Regarding the effect of the pauses, we did not find significant evidence that they enhance the segmentation —there was a marginally significant effect on the pupil dilation, but no a significant interaction, thus it was probably result of between subjects variability. Despite the lack of an effect of familiarization, the patter of results was in the direction expected if pauses help the segmentation: pupil dilation for the group in the pause condition showed the patter s > p > w, whereas in the continuous condition s > p $\approx$ w; and the duration of the first fixation for Switch-words in the pause condition was the only one that appeared a bit longer.

Finally, when we looked to individual data to see if infants performing better were showing longer looking times for either Part-words or Switch-words, we did not find any positive result, not allowing as to make conclusions about the representation of the sequence infants create after segmentation.

All together our results suggest that 5-month-old infants are sensitive to distributional cues over tones, but presumably a longer familiarizations phase than for linguistic stimuli is required for them to extract the tones sequences.

5.8 Chapter Discussions

In the current chapter we investigated how adults and infants encode information extracted from a flow of stimuli. We did it across multiple domains, and we explored if it differs between a condition in which only distributional cues can be used to chunk the stream and another in which also brief brakes are inserted.

The motivation of these experiments was to see if a more abstract representation of the segmented sequences, in particular the association of each item constituting the sequence to given a position, was created; or if only local associations between the constituting items were retained. The experiments with adults suggest that in the case of speech the representation of the segmented words relies more on the position of the syllables. The creation of this representation seems not to depend of the type of cues signalling the edges —even if the pauses do facilitate the segmentation— but instead it seems to depend of how well the words were extracted: once participants get the words, independently of how, the representation is not any more based only on local associations.

The results for both visual experiments were very different than for speech. Participants seem to manly retain the association between pairs of items and not associate them to a position in a sequence. This was independent of how well they had learned the distributional properties of the stream and independent of the presence of other cues (brief brakes) signalling the edges, even if this cues do enhance adult performance.

Our results show important differences in the processing of speech and visual sequences, but the visual and auditory modalities are very different in terms of the type of stimuli they involved. In the auditory modality information is basically only integrated along one dimension, time. Meanwhile in the visual modality information has to be integrated along both space and time. This implies that patterns in auditory stimuli are necessarily identified along time, which is not the case for visual stimuli. Therefore, the contrast between Experiment 6 and Experiments 8 and 9 could be either consequence of this difference between modalities or it could be specific to speech processing. Unfortunately we cannot answer this question because adults performance in segmenting the stream of tones was very low. Observed that the effect is not strictly language specific because we did not observe it when syllables were presented on the screen.

Two fundamental questions arise. First, is this more abstract encoding of positional information in speech an innate bias or does it arise as a consequence of adults experience with language? Second, is this effect specific to the auditory modality or to processing speech?

In Experiment 10 we aimed to answer the first question and we tested 5-month-old, who have much reduced experience with language, on the exact same task than adults in Experiment 6. To do so we used pupil dilation as a measure of surprise. Infants succeed in the segmentation task, which constitutes the first behavioural evidence of this ability in such young infants. Nevertheless, we obtained the same level of pupil dilation toward Part-words and Switch-words, which did not allow us to make conclusions regarding how they encode the words. As we proposed before, a simpler familiarization stream could be used to try to obtain a different response for the two word classes. Another thing should be checked as well in order to stablish how infants encode information

extracted from continuous speech. It may be the case that 5-month-old do not consider the synthesized speech as real language, hence they do not encode the extracted chunks in the same way than they would encode words from speech. In order to control for this, an experiment using more ecological stimuli should be done. In sum, if a more abstract encoding of speech is acquired or innate remains as an open question.

In our experiments neither adults nor infants succeeded to segment the stream of tones, leaving the second question regarding the domain specificity of the effect open as well. As we have already extensively discussed, this failure to extract the tones sequences questions previous finding showing the same level of performance in the segmentation of speech and non linguistic auditory stimuli.

Despite both questions remain to be answered in future work, our results have clear consequences in the understanding of language processing. An abstract encoding is particularly useful in a hierarchical system like language, in which positional information is fundamental, and we demonstrated that this abstract encoding occurs in speech, and moreover it occurs even when no other cues, besides the distribution of syllables, that signal the edges. As we pointed in the discussion of Experiment 6, this contradicts Peña and colleagues (Peña et al., 2002) hypothesis. If statistical and rule learning are two independent mechanisms, or if they are part of the same leaning mechanism (Christiansen et al., 2006) cannot be said from our results, nevertheless, what our experiments do suggest, is that rule based representations can be created even when the only available information are distributional cues.

The results also open many questions regarding the bases of positional encoding of speech. We tested a very specific type of sequences, in which the first syllable of a word was replaced by the first syllable of another word. To be certain that the observed effect is related with the association of syllables to a given position and to understand its extension, other type of Switch-words should be tested. It would be particularly interesting to see if this applies only to the first position or if it extends to other positions. Note that some care has to be taken to test switches between middle syllables. If for example the middle syllable in the Words of our experiments are exchanged, the obtained sequence has dependencies between syllables that belie even more the distributional properties of the familiarization. An appropriate type of Part-word should be used for this comparison.

To finalize, a methodological remark. We implemented a new methodology for testing speech segmentation in infants. Our protocol enables to record two independent behavioural measures: looking times and pupil dilation. Looking times present the problem that whether novelty or familiarity effect will be observed is difficult to predict (Hunter, Michael A; Ames, 1988), making hard the interpretation of results. Pupil dilation is a much more automatic response and the directionality of the effects is highly predictable. Moreover, our results bring

evidence that the two measures do not reflect the same, thus their combined use opens new possibilities for the study of infant cognition.

Chapter 6

General Discussion

The aim of this thesis was the study of the mechanism available from birth for segmenting speech —in particular statistical learning and prosody—, of memory constraints for the encoding of sequential information, and of how these elements interact for bootstrapping linguistic information. We also investigated the neural bases of speech perception in neonates. The results of the individual studies have been discussed in their corresponding chapters. In this final chapter, I present a summary of the most relevant findings together with some questions that remain open, and I draw some general conclusions and present some speculations. To present them in a concise way I identify three main themes. First, findings regarding the learning mechanism and cognitive biases that can play a role in parsing speech. Second, understanding the functional architecture of the human brain at birth. And third, methodological contributions to the study of infant cognition.

6.1 Parsing speech: Learning mechanism and cognitive biases

In Chapter 2 we showed that neonates can extract words from continuous speech using both the distribution of syllables and prosodic information. Previous researches have shown neonates sensitivity to distributional information in the speech signal (Teinonen et al., 2009), and to well formed prosodic units (Christophe et al., 1994), but our results go beyond. They provide evidence that newborns extract information out of it and form some kind of representation. Importantly, this representation appears to overlook some properties of the stimulus, as it was clearly demonstrated in Experiment 2. The results agree with previous studies demonstrating that infants at birth can create short term memories of words (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012) and they neglect voice differences in the recognition of phonemes (Jusczyl, Pisoni, and Mullennix, 1992; Kuhl, 1983; Mahmoudzadeh et al., 2013). These findings suggest the existence of a phonological loop or a prototype of it since birth.

These results pose some questions as well.

The capacity of learning from distributional cues is available for different type of stimuli since birth (Kudo et al., 2011; Bulf, Johnson, and Valenza, 2011). Experiment 1 demonstrates extraordinary neonates' abilities for extracting words based on distributional information. Is the performance equal in non linguistic domains? or are infants better in speech processing?

Experiment 2 shows that neonates can recognize prosodic units, but which is the origin of this knowledge? Is it because of the existence of perceptual biases or a consequence of prenatal linguistic experience?

Finally, we can ask what happens when distributional and prosodic cues are present together. Do neonates rely more on one cue than on the other?

Experiments in Chapter 4 demonstrate that even from birth not all the elements of a sequence are encoded in an equal way. As in adults (Henson, 1998; Hurlstone, Hitch, and Baddeley, 2014), the first and last elements are better encoded. Remarkably, in Experiment 5, we found that a prosodic cue —a subtle pause— affects the encoding, specifically it enhances the encoding of intermediate elements. These results suggest that the memory system of infants at birth is governed by similar constraints as the memory system in adults; but at the same time they reveal the existence of a mechanism available since early in life for the encoding of long sequences, a fundamental requirement for language acquisition. In particular, the study offers evidence for a critical role of even subtle prosodic cues in speech processing since birth.

Our results suggest that similar principles operate for the encoding of sequential information in neonates and adults. Nevertheless, further investigation is needed.

First, it is not completely clear from our experiments if neonates are encoding position relative to the edges as adults apparently do; or if neonates are better encoding either the first, or the last, or both syllables.

Second, how does the pause affect the encoding? Our interpretation is that it acts as a segmentation cue—it generates an edge— meaning that infants encode two separated units. Another explanation could be that the sequence is still stored as a unique unit but that there is a switch of attention to intermediate elements. A complete understanding of how the representations generated from sequential informations are affected by prosodic cues needs more investigation in both infants and adults.

In Chapter 5 we observed that when adults segment speech they do not only encode the co-occurrence of syllables. They appear to form a more abstract representation of the segmented units by associating the syllables with their position in the sequence. In particular they prefer a sequence in which the syllables occupy the same position as in real words even if are new to them, than a sequences they have heard but in which syllables are shifted. Remarkably, we did not observe

similar effect in the visual modality, not even when linguistic stimuli were used. The results reveal that adults process speech differently than information presented sequentially in the visual modality. Unfortunately, because adults failed at segmenting auditory non linguistic information, we cannot rule out that the effect is characteristic of the processing of sequential information in the auditory modality. We believe that an encoding system based on positions sets basis for abstract rule learning in language. For example adults may have encoded that in Italian, articles usually appear at the beginning of phrases.

An interesting aspect of our findings is that the generation of an abstract representation is independent of the type of segmentation cues available. Pauses between words in the stream helped the segmentation as expected, but the encoding of position does not appear to depend of its presence; instead it appears to occur as soon as the units are identified independently of how. This challenges the idea that rule learning occurs only when the information is presented segmented (Peña et al., 2002), and suggests that rules can be extracted from pure distributional information.

In order to explore if this particular processing of linguistic information observed in adults was the result of their extended experience with language, or was due to biological constraints, we tested 5-month-old infants. However our results did not provide evidence of an abstract encoding and we cannot make strong conclusions because we did not obtain negative evidence either; exactly when infants start to generate this abstract encoding remains an open question.

Our combined results show that since birth infants are endowed with a cognitive system ready for the processing of complex sequential information as speech. Parsing speech probably relies on the combined used of multiple cues present in the input, and in the restrictions imposed by the cognitive system. We have demonstrated that some perceptual biases and memory constraints that operate for adults during speech processing, exist since birth. But at the same time, results from Experiments 6 and 10 suggest that experience, together with innate constraints may give place to new biases in processing.

Distributional cues are fundamental in language; however, how distributional cues are processed is intricate. First, it is susceptible to perceptual biases, in particular we described perceptual biases operating on prosodic cues. Second, it is affected by constraints imposed by the memory system. And third, these perceptual biases and memory constraints differ across domains and may change during development. In other words, statistical learning means only the extraction of regularities from the input based on the co-occurrence of items; but, how are the items defined and how are they encoded? Differences on these aspects would lead to distinct learning processes between different types of stimuli, even if with similar distributional properties.

Moreover, not only qualitative but also quantitative differences, should be investigated. For example, are infants better for processing and encoding linguistic information compared with other kinds of stimuli? Statistical learning performance across domains and during development should be explored, including the investigation of the type of information extracted.

In brief, language learning surely relies on the computation of distributional cues —after all distributional cues are a fundamental difference between languages. Nevertheless, infants may be gifted with a neural system particularly adapted for the processing of regularities and the encoding of speech.

We believe the results reported in this thesis, open new perspectives to the study of language acquisition in particularly regarding parsing. We think that the distinction of learning mechanisms (statistical learning, rule learning, perceptual biases) should be done more carefully, and that the combined used of multiple sources of information should be further investigated.

We also demonstrated a salient role of prosody in parsing since very early in life. We believe that prosody could restrict the space, otherwise huge, on which statistical learning operates, and could play a fundamental role in the extraction, encoding and processing of speech structure. Prosody appears as a good candidate because of different reasons. Prosodic units are easily detected since birth (Hirsh-Pasek et al., 1987; Nelson et al., 1989; Gerken, Jusczyk, and Mandel, 1994; Nazzi, Jusczyk, and Johnson, 2000; Soderstrom et al., 2003; Christophe et al., 1994), either due to perceptual biases or to exposure to the language; and it has different components that operate at different levels of the speech (Hayes, 1989; Nespor and Vogel, 2007; Langus et al., 2012). Moreover it provides a substrate on which language specific mechanisms can work.

We propose that prosody, together with an abstract encoding of positional information, may provide a base for the processing and acquisition of speech structure. The idea is as follows:

- Prosodic units mark edges.
- Different edges can be marked simultaneously by distinct prosodic components.
- The position of the different components of speech —syllables, words, and even phonological phrases— is encoded relative to the edges of the prosodic unit in which they are contained.
- Statistical regularities are identified mainly inside prosodic units. Moreover, statistical regularities are not only based on the co-occurrence of units but also on how often a given element appears in a determined position.

A mechanism like this could settle the bases for a hierarchical organization, or at least for a parallel processing in language (Frank, Bod, and Christiansen, 2012). Of course our results are far from validate this hypothesis, but they do demonstrate that many of these fundamental elements are present since very early in life.

6.2 Neonates' brain functional organization

In all the fNIRS studies we performed we observed consistent activation over frontal areas in response to novel stimuli, denoting that even at birth associative cortices are active. These results add to recent findings showing activation over frontal areas in similar tasks (Benavides-Varela et al., 2011; Benavides-Varela et al., 2012; Mahmoudzadeh et al., 2013; Nakano et al., 2009), and contradicts the classical view that frontal areas are not functionally active early in life (see for example for a review Johnson, 2001; Dehaene-Lambertz and Spelke, 2015). The effect we observed was broad and extended mainly over prefrontal and temporal regions. Further investigation is needed to understand the origin and implications of this habituation/ novelty effect.

In Chapter 3 we performed a dynamic functional connectivity analysis, that provided different interesting results.

First, the main properties of the functional architecture were not dependent on the task. In particular, homologous regions and close areas were functionally related. Moreover, we identified stable and reproducible functional networks with small-world properties, denoting a functional architecture with segregated process that notwithstanding shares information, present since birth. Moreover, these networks show connections with opposite behaviours: an increase in connectivity between some areas of the brain was associated with a decrease in other areas. Specifically, we observed an antagonism between frontal vs. more posterior areas, left hemisphere vs. right hemisphere, and connections involving parietal areas vs. more fronto-temporal areas.

Second, the dynamics of the functional connectivity had periods of high global and low global connectivity. Remarkably, this variation in the global connectivity was task dependent, with stronger values when neonates listened to speech with structure (build by concatenating words). The global strength of the connections seems an indicator of infant's cognitive state, encouraging its study.

Third, the extraction of the words from continuous speech appears to require the integration of information between distant areas within hemispheres.

Fourth, our results suggest a left hemisphere dominance to process distributional cues during segmentation, and a right dominance for the processing of supra-segmental information —prosody. This left dominance goes along with faster maturation of the left arcuate fasciculus.

Overall our results indicate that functional activity has a precise organization since birth, with stable properties across time and tasks. For future work it would be interesting to investigate to which extent the functional networks identified reflect direct anatomical connections. Despite the appearance of a stable functional organization, we did find a modulation of the activity by infant's cognitive state

and specifically by the stimuli presented —while hearing a stream with informative transitional probabilities, random syllables, and during resting state. These results offer evidence of a highly reactive neural system since birth. Moreover the results were consistent with anatomical studies describing brain maturation in the first trimester of life, creating a link between structure and cognition.

Finally, I would like to remark that we explored for the first time a new dimension of the functional architecture of infants' brain: its changes along time. The results invite to study the dynamical aspects of functional connectivity in infants. The origin of the observed transient periods of high connectivity is not clear and its relation with the cognitive state appears a promising field for future research.

6.3 Methodological contributions

A special effort was put on different methodological aspects, for both fNIRS and eye-tracking experiments.

1. The pre-processing protocol for the fNIRS data applied in Experiments 1 and 2 was optimized in particular for the identification and corrections of motion artifacts.

2. In all the fNIRS studies we implemented an habituation protocol. The protocol takes advantage of the strong habituation effects that are present in fNIRS data. Even if the origin of the effect is not clear, not allowing to make strong conclusion regarding the neural bases of the tested hypothesis, the method certainly enables us to evaluate neonates' cognitive abilities. To obtain behavioural responses from neonates is hard, and because fNIRS is a non invasive method, habituation protocols using fNIRS can be used even with this strict purpose.

3. To our knowledge, we implemented the first dynamic functional connectivity analysis based on fNIRS recordings, and the first dynamic functional analysis on infants data. The two analyses seem to be suitable for fNIRS, and in particular the PCA analysis allows to identify functional networks based on the temporal evolution of the connections. Our results encourage the use of fNIRS for the study of functional connectivity.

4. In Experiments 10 and 11, we implemented a new protocol that uses pupil dilation as dependent measure to test speech segmentation in infants. Pupil dilation is an automatic measure, for which the directionality of the effects is easy to predict. The protocol represents an alternative to the traditional head-turning methodology base on looking times.

Appendix A

Appendix to Chapter 2

A.1 Appendix to Experiment 1

The results using *Hb* for both the cluster based permutation analysis and mean activation analysis are presented below.

Cluster Based Permutation Analysis.

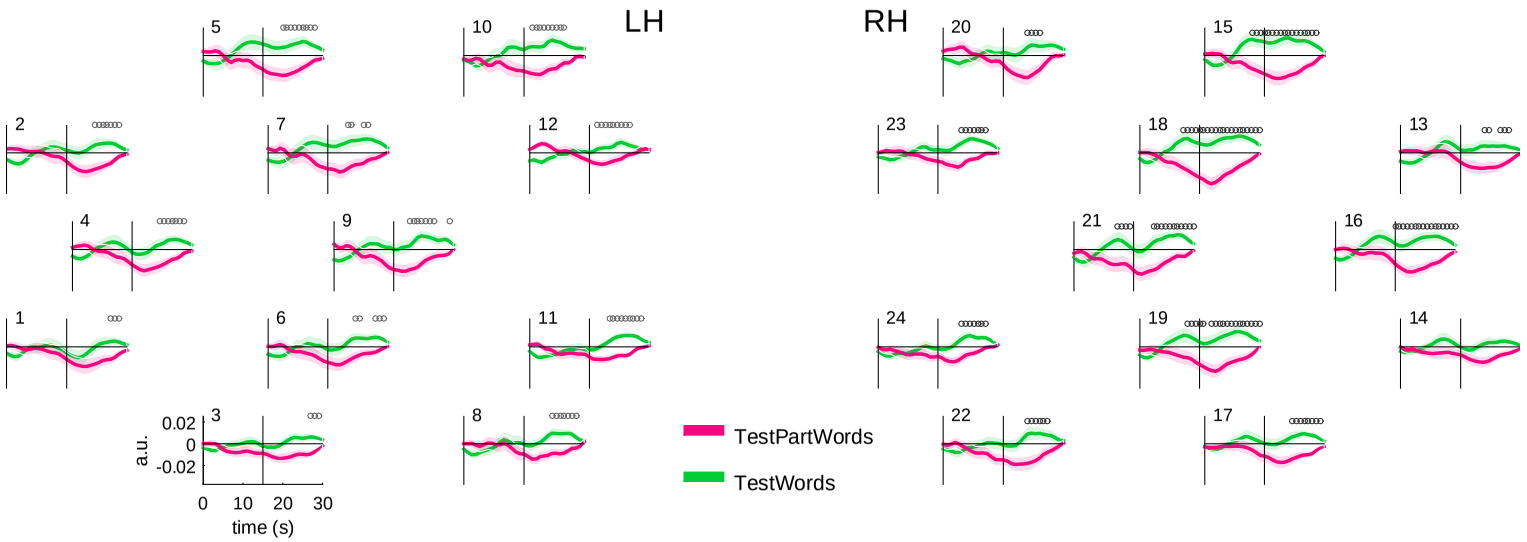


FIGURE A.1: Cluster based permutation analysis for Experiment 1 using Hb. HRFs for Words (green) and Part-words (pink) during test blocks. Vertical lines indicate the onset and offset of the stimulus. Marks below the individual channels show the time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

Mean Activation Analysis.

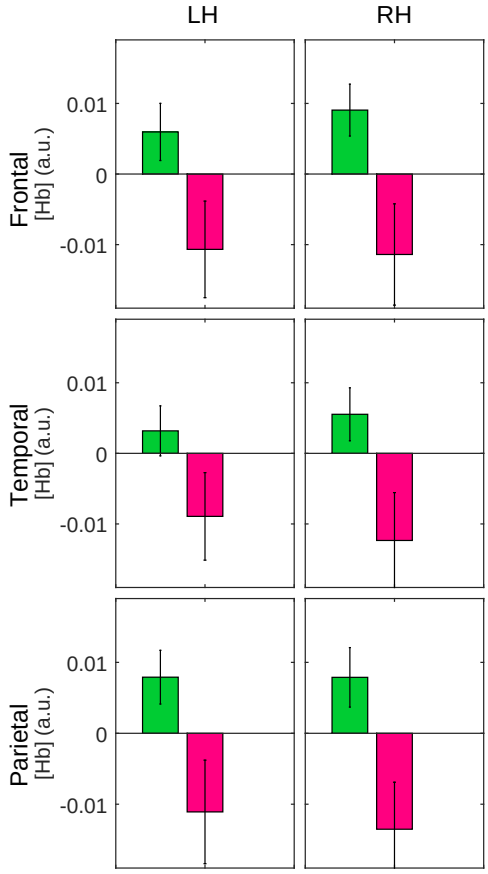


FIGURE A.2: Mean activation analysis for Experiment 1 using Hb. The mean activity for Words (green) and Part-words (pink) over each of the three regions of interest in the left and right hemisphere is shown. Error bars represent standard errors

Hb	SumSq	DF	MeanSq	F	P
Hemisphere	0.28030	39	0.00719		
Error (Hemisphere)	0.00000	1	0.00000	0.01983	0.88870
Region	0.00878	39	0.00023		
Error (Region)	0.00016	2	0.00008	0.37840	0.68620
Condition	0.01623	78	0.00021		
Error (Consdition)	0.03852	1	0.03852	6.62194	0.01400
Hemisphere:Region	0.22687	39	0.00582		
Error (Hemisphere: Region)	0.00012	2	0.00006	0.55470	0.57650
Hemisphere:Condition	0.00872	78	0.00011		
Error (Hemisphere: Condition)	0.00048	1	0.00048	2.26312	0.14050
Region:Condition	0.00827	39	0.00021		
Error (Region: Condition)	0.00057	2	0.00028	1.29205	0.28050
Hemisphere: Region: Condition	0.01713	78	0.00022		

TABLE A.1: Statistical analysis for Experiment 1. 3-ways ANOVA using Hb as dependent measure. Hemisphere (left/ right), region (frontal/ temporal/ parietal) and condition (Edge/ Internal) are within subject factors.

A.2 Appendix to Experiment 2

The results using Hb for both the cluster based permutation analysis and mean activation analysis are presented below.

Cluster Based Permutation Analysis.

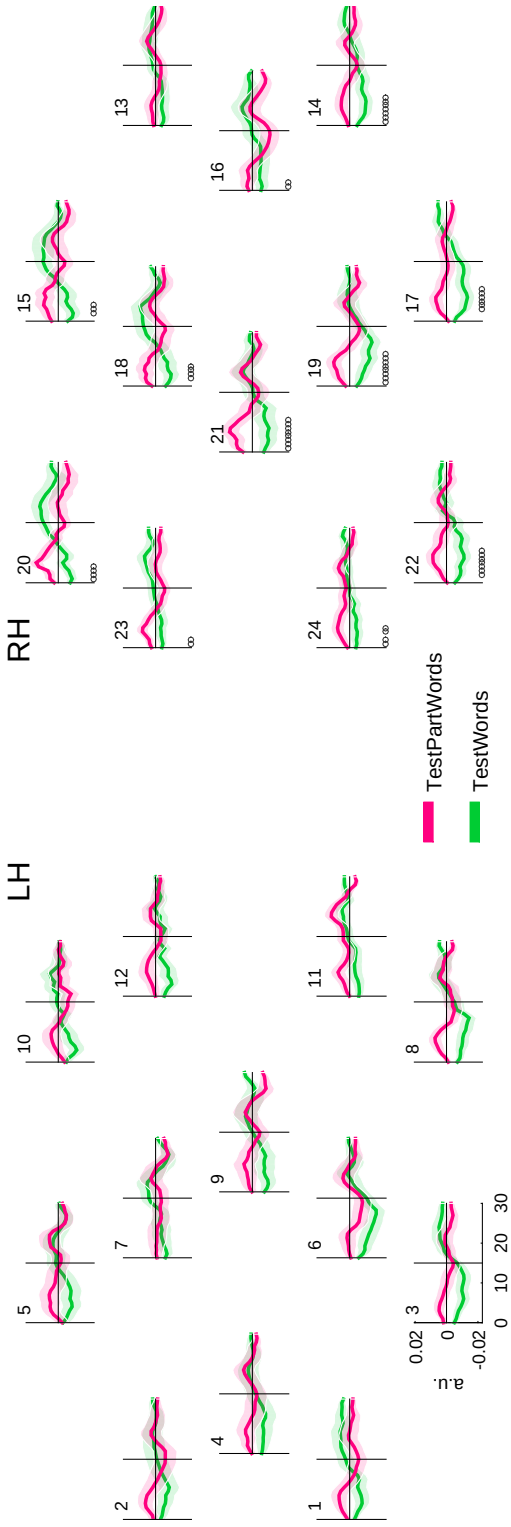


FIGURE A.3: Cluster based permutation analysis for Experiment 2 using *Hb*. HRFs for Words (green) and Part-words (pink) during test blocks. Vertical lines indicate the onset and offset of the stimulus. Marks below the individual channels show the time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

Mean Activation Analysis.

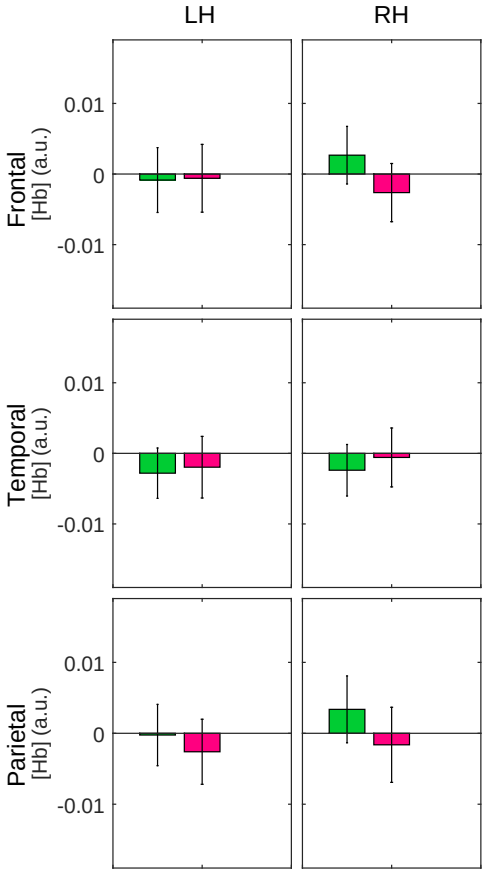


FIGURE A.4: Mean activation analysis for Experiment 2 using *Hb*. The mean activity for Words (green) and Part-words (pink) over each of the three regions of interest in the left and right hemisphere is shown. Error bars represent standard errors

Hb	SumSq	DF	MeanSq	F	P
Hemisphere	0.16651	39	0.00427		
Error (Hemisphere)	0.00021	1	0.00021	2.07686	0.15750
Region	0.00390	39	0.00010		
Error (Region)	0.00028	2	0.00014	0.86375	0.42560
Condition	0.01268	78	0.00016		
Error (Condition)	0.00032	1	0.00032	0.08990	0.76590
Hemisphere: Region	0.13761	39	0.00353		
Error (Hemisphere: Region)	0.00006	2	0.00003	0.24048	0.78680
Hemisphere: Condition	0.00951	78	0.00012		
Error (Hemisphere: Condition)	0.00017	1	0.00017	1.15619	0.28890
Region: Condition	0.00588	39	0.00015		
Error (Region: Condition)	0.00055	2	0.00028	1.49472	0.23070
Hemisphere: Region: Condition	0.01437	78	0.00018	1.00000	0.50000

TABLE A.2: Statistical analysis for Experiment 2. 3-ways ANOVA using *Hb* as dependent measure. Hemisphere (left/ right), region (frontal/ temporal/ parietal) and condition (Edge/ Internal) are within subject factors.

Appendix B

Appendix to Chapter 3

B.1 Appendix to the connectivity analysis for Experiment 1

The eigen-networks 5 to 8 are represented bellow.

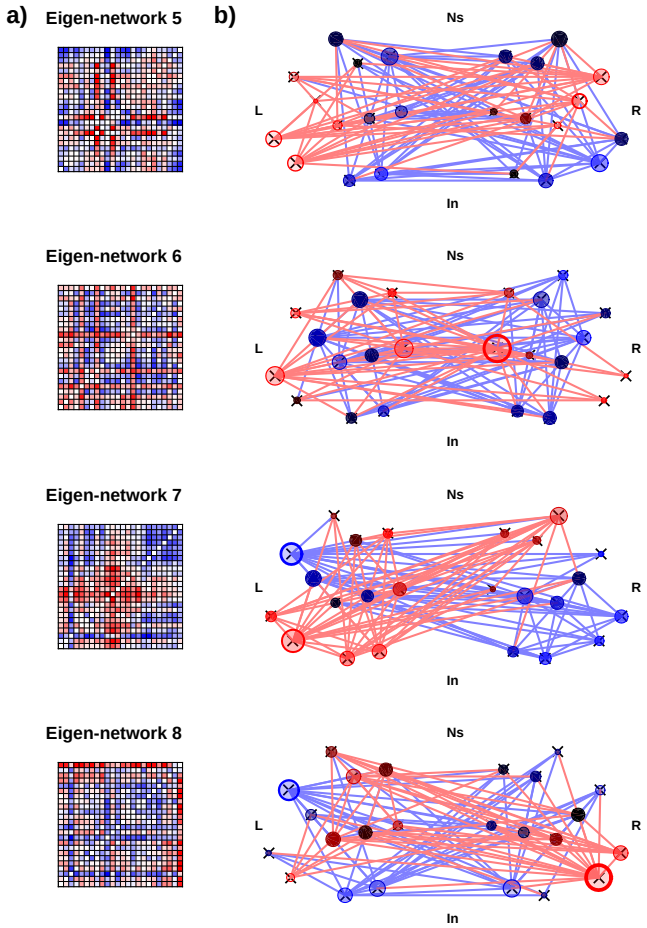
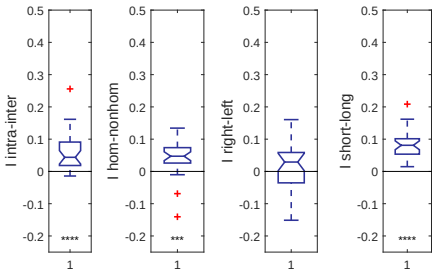
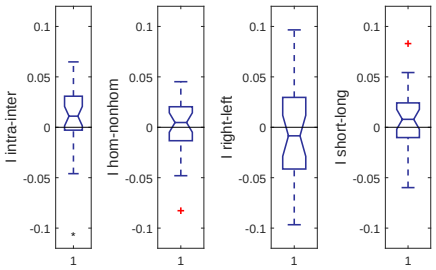


FIGURE B.1: Eigen-networks 5 to 8 for Experiment 1 using HbO_2. The representation uses the same code than in in previous figures.

Strength and variability of the connections. Hb



(A) Strength

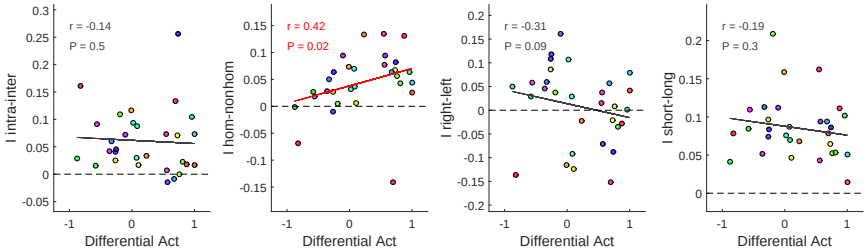


(B) Variability

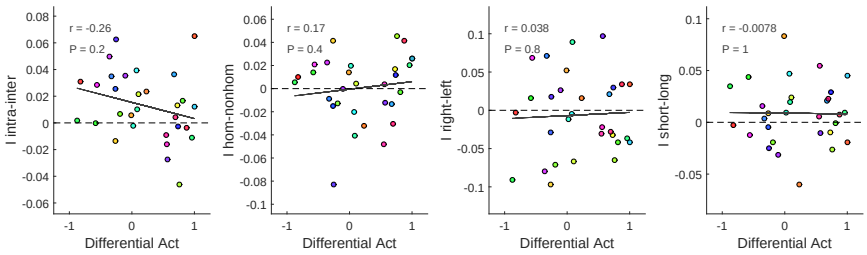
FIGURE B.2: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 1. Connectivity was estimated using Hb. Asterisks bellow represent Bonferroni corrected P-values of the t-test against zero.

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength	I intra-inter	0.0607	0.0573	0.0393	0.0821	29	5.8019	0.0000	0.0000
Hb	I hom-nonhom	0.0442	0.0563	0.0232	0.0652	29	4.3066	0.0002	0.0010
	I rigth-left	0.0081	0.0810	-0.0222	0.0383	29	0.5470	0.5886	2.3540
	I short-long	0.0855	0.0396	0.0707	0.1003	29	11.8103	0.0000	0.0000
Variability	I intra-inter	0.0130	0.0253	0.0036	0.0225	29	2.8160	0.0087	0.0350
Hb	I hom-nonhom	0.0008	0.0284	-0.0099	0.0114	29	0.1486	0.8829	3.5320
	I rigth-left	-0.0061	0.0528	-0.0258	0.0136	29	-0.6319	0.5324	2.1300
	I short-long	0.0087	0.0295	-0.0023	0.0197	29	1.6130	0.1176	0.4700

TABLE B.1: Statistical analysis on the different indexes For Experiment 1 using Hb. T-tests against chance. Bonferroni correction was used.



(A) Strength



(B) Variability

FIGURE B.3: Spearman correlations between task performance and functional connectivity for Experiment 1 using Hbs. The differential activation for part-words and words during test blocks (x-axis) is plotted against the different indexes estimated during the familiarization phase (y-axis): $I_{inter-intra}$, $I_{hom-nonhom}$, $I_{right-left}$, $I_{short-long}$. Each dot represent a subject.

.

Principal Components Analysis. Hb

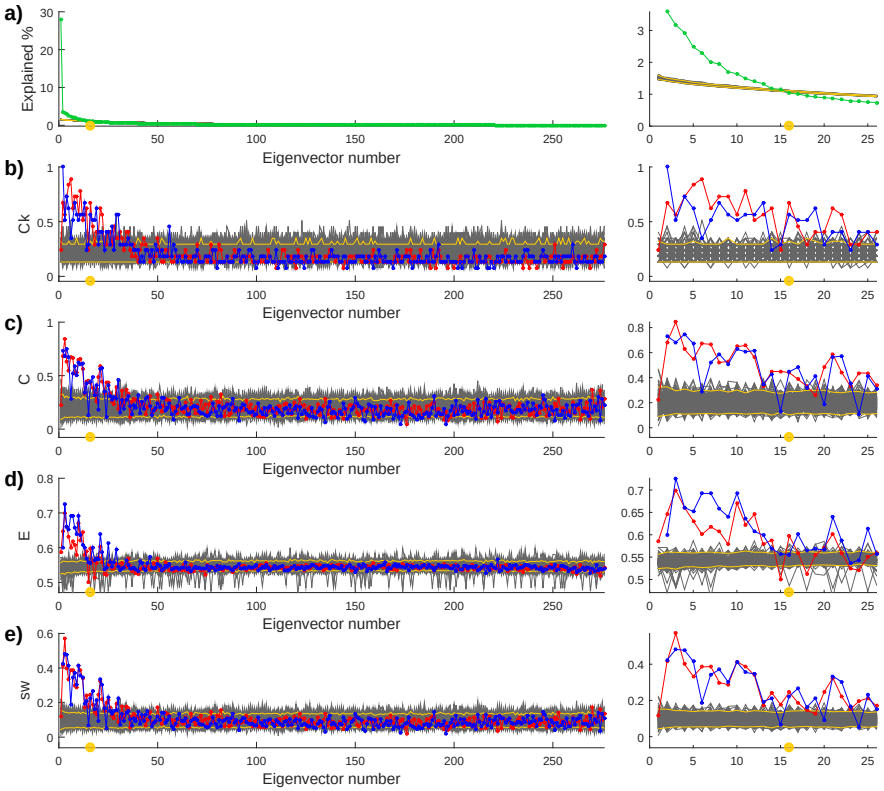


FIGURE B.4: Phase randomization simulations for Experiment 1 using Hb. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

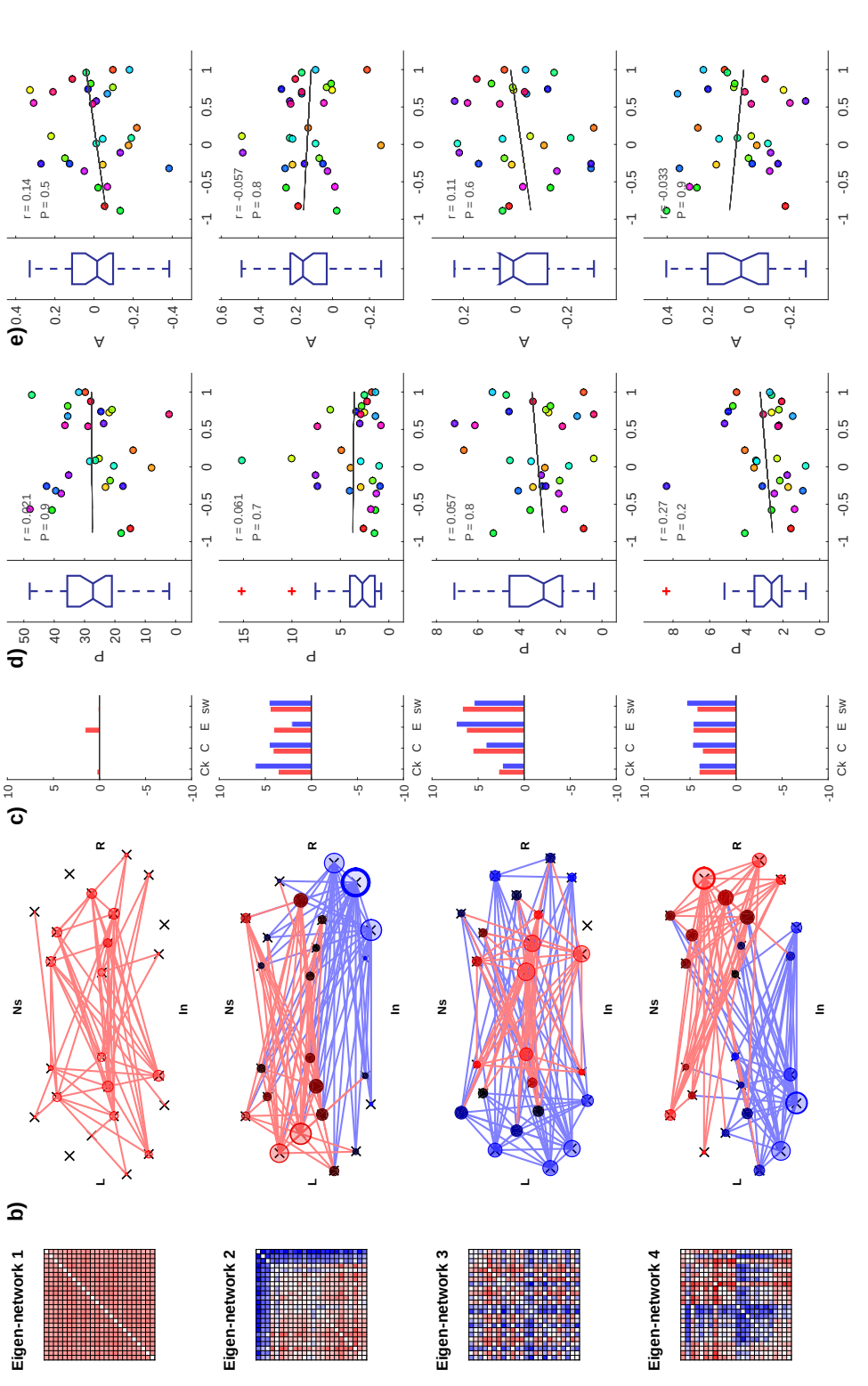


FIGURE B.5: PCA analysis for Experiment 1 using HbO. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the with of its line the centrality of the node. **c)** Z-scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Spearman correlations between the explained variability (P) for each eigen-network and the differential activation for Part-words and Words. **e)** Spearman correlations between the asymmetry (A), for each eigen-network and the differential activation for Part-words and Words.

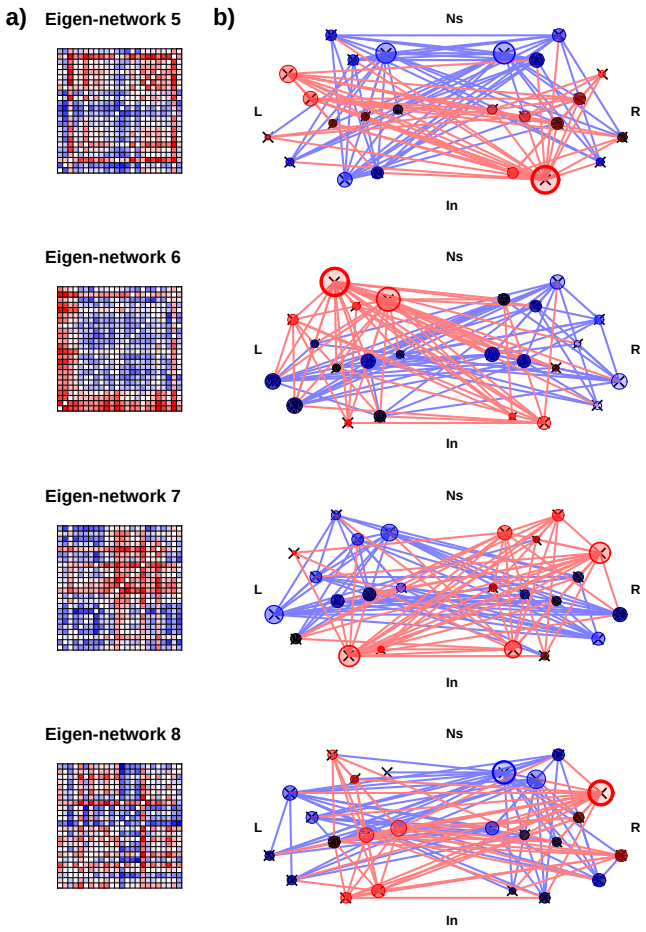


FIGURE B.6: Eigen-networks 5 to 8 for Experiment 1 using Hb. The representation uses the same code than in previous figures.

B.2 Appendix to the connectivity analysis for Experiment 3

The eigen-networks 5 to 8 are represented bellow.

Strength and variability of the connections. Hb

		SumSq	DF	MeanSq	F	P
Strength Hb	Condition	0.571	2	0.285	9.05537	0.00050
	Error (Condition)	1.386	44	0.032		
Variability Hb	Condition	0.074	2	0.037	4.31324	0.01950
	Error (Condition)	0.378	44	0.009		

TABLE B.2: Statistical analysis on the strength of the connections for Experiment 3 using Hb. 1-way ANOVA with condition as within subject factor.

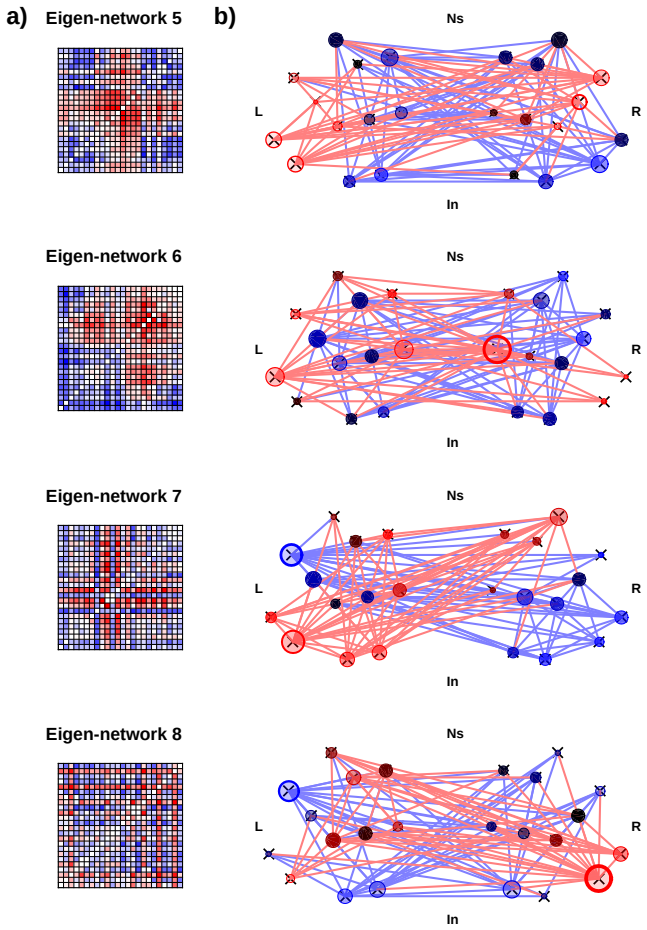
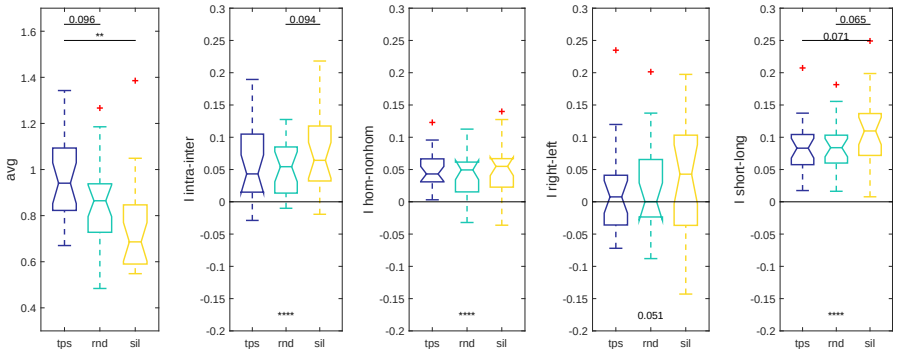


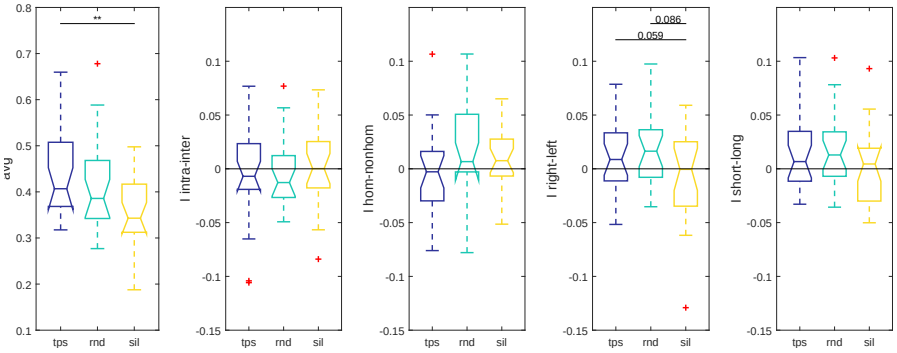
FIGURE B.7: Eigen-networks 5 to 8 for Experiment 3 using HbO_2. The representation uses the same code than in previous figures.

		Difference	StdErr	P	CI lower	CI upper
Strength	TPs-Rnd	0.112	0.051	0.09576	-0.017	0.240
Hb	TPs-Sil	0.223	0.054	0.00114	0.088	0.360
	Rnd-Sil	0.111	0.052	0.10889	-0.021	0.240
Variability	TPs-Rnd	0.021	0.030	0.76220	-0.054	0.100
Hb	TPs-Sil	0.078	0.022	0.00564	0.022	0.130
	Rnd-Sil	0.057	0.029	0.15303	-0.017	0.130

TABLE B.3: Statistical analysis on the strength of the connections for Experiment 3 using Hb. Post-hoc multiple comparisons analysis between conditions. Turkey-Kramer correction was used.



(A) Strength



(B) Variability

FIGURE B.8: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 3. Connectivity was estimated using Hb. The for each of the conditions is presented: TPs (blue), Rnd (green) and Sil (yellow). Asterisks bellow represent Bonferroni corrected P-values of the t-test against zero. Asterisks above show the results of the post-hoc multiple comparisons analysis using Turkey-Kramer correction.

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength Hb	I intra-inter	0.0624	0.0550	0.0492	0.0756	68	9.4242	0.0000	0.0000
	I hom-nonhom	0.0475	0.0368	0.0387	0.0564	68	10.7210	0.0000	0.0000
	I rigth-left	0.0239	0.0776	0.0052	0.0425	68	2.5554	0.0128	0.0510
	I short-long	0.0932	0.0451	0.0824	0.1041	68	17.1512	0.0000	0.0000
Variability Hb	I intra-inter	-0.0035	0.0371	-0.0124	0.0054	68	-0.7825	0.4366	1.7470
	I hom-nonhom	0.0081	0.0380	-0.0010	0.0173	68	1.7795	0.0796	0.3190
	I rigth-left	0.0067	0.0402	-0.0030	0.0163	68	1.3774	0.1729	0.6920
	I short-long	0.0093	0.0341	0.0011	0.0175	68	2.2672	0.0266	0.1060

TABLE B.4: Statistical analysis on the different indexes For Experiment 1 using Hb. T-tests against chance. Bonferroni correction was used.

		SumSq	DF	MeanSq	F	P
Strength: $I_{intra-inter}$ Hb	Condition	0.008	2	0.004	2.58437	0.08690
	Error (Condition)	0.069	44	0.002		
Strength: $I_{hom-nonhom}$ Hb	Condition	0.002	2	0.001	1.01052	0.37230
	Error (Condition)	0.048	44	0.001		
Strength: $I_{right-left}$ Hb	Condition	0.004	2	0.002	0.64218	0.53100
	Error (Condition)	0.136	44	0.003		
Strength: $I_{short-long}$ Hb	Condition	0.008	2	0.004	4.26064	0.02040
	Error (Condition)	0.041	44	0.001		
Variability: $I_{intra-inter}$ Hb	Condition	0.001	2	0.000	0.45157	0.63950
	Error (Condition)	0.048	44	0.001		
Variability: $I_{hom-nonhom}$ Hb	Condition	0.007	2	0.003	2.23637	0.11890
	Error (Condition)	0.066	44	0.002		
Variability: $I_{right-left}$ Hb	Condition	0.011	2	0.005	3.69288	0.03290
	Error (Condition)	0.063	44	0.001		
Variability: $I_{short-long}$ Hb	Condition	0.003	2	0.001	0.87997	0.42200
	Error (Condition)	0.064	44	0.001		

TABLE B.5: Statistical analysis on the different indexes for Experiment 3 using Hb. 1-way ANOVA with condition as within subject factor.

Principal Components Analysis. Hb

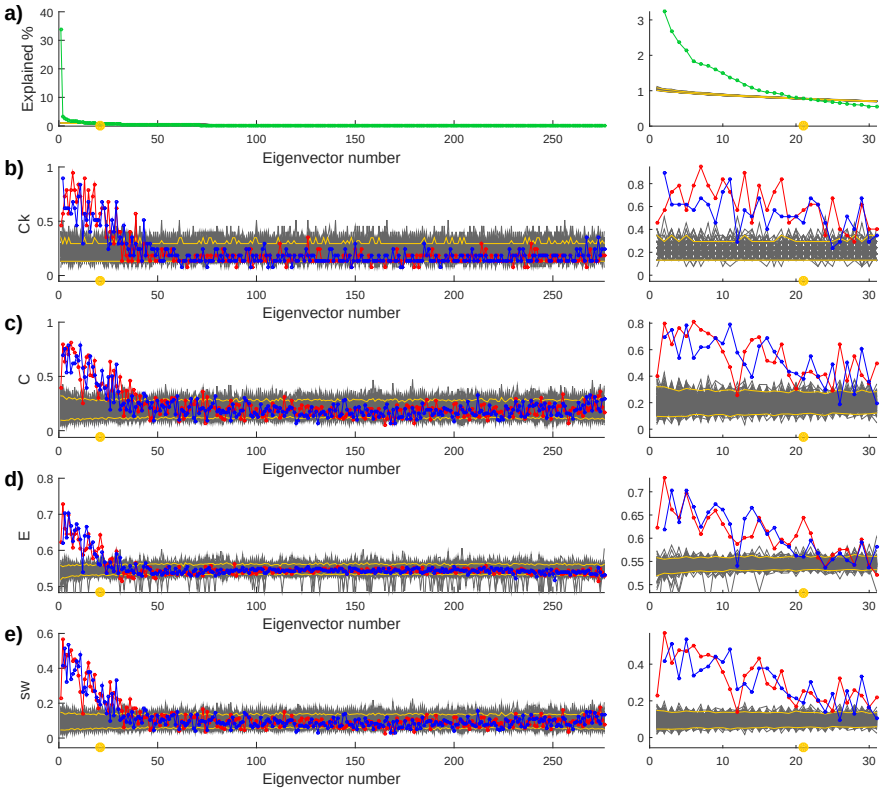


FIGURE B.9: Phase randomization simulations for Experiment 3 using Hb. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

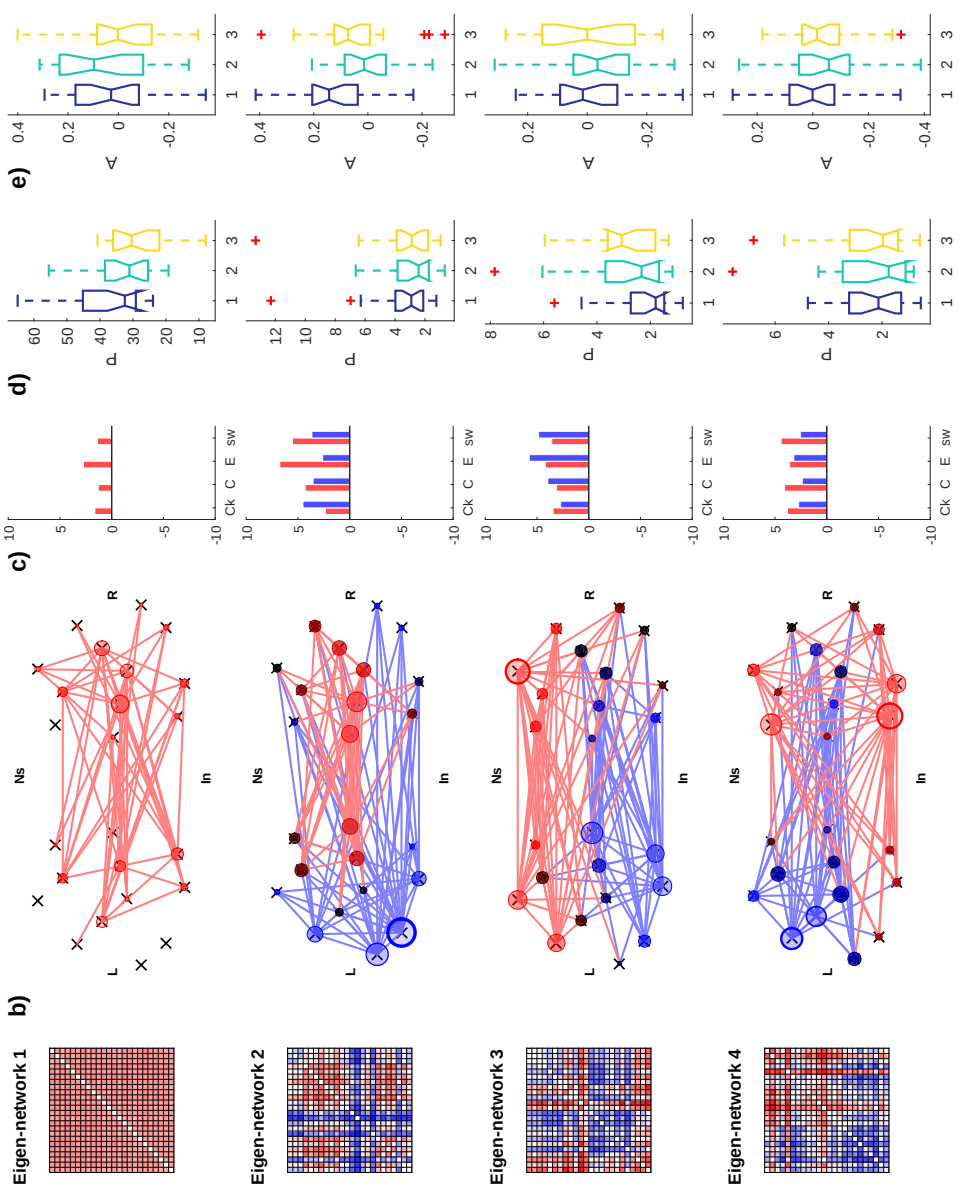


FIGURE B.10: PCA analysis for Experiment 3 using Hb. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the with of its line the centrality of the node. **c)** Z-scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Explained variability (P) for each eigen-network per condition. TPs (blue), Rnd (green) and Sil (yellow). **e)** Asymmetry (A) per condition.

		SumSq	DF	MeanSq	F	P
P	Condition	165.0	2	82.5	4.3	0.02040
Hb	Error (Condition)	852.1	44	19.4		
	Eigen-network	46989.4	3	15663.1	529.1	0.00000
	Error (Eigen-network)	1953.8	66	29.6		
	Condition: Eigen-network	613.9	6	102.3	3.6	0.00240
	Error (Condition: Eigen-network)	3737.1	132	28.3		
A	Condition	0.131	2	0.06546	2.347	0.10750
Hb	Error (Condition)	1.227	44	0.028		
	Eigen-network	0.368	3	0.12260	4.608	0.00550
	Error (Eigen-network)	1.756	66	0.027		
	Condition: Eigen-network	0.157	6	0.026	1.128	0.34960
	Error (Condition: Eigen-network)	3.060	132	0.023		

TABLE B.6: Statistical analysis on the explained variance and asymmetry for Experiment 3 using *Hb*. 2-way ANOVA with with condition and eigen-network as within subject factors.

		Difference	StdErr	P	CI lower	CI upper
P	TPs-Sil	1.889	0.588	0.01078	0.411	3.370
Hb	TPs-Rnd	0.831	0.704	0.47697	-0.937	2.600
	Rnd-Sil	1.059	0.649	0.25419	-0.572	2.690

TABLE B.7: Statistical analysis on the explained variance for Experiment 3 using *Hb*. Post-hoc multiple comparisons analysis between conditions. Turkey-Kramer correction was used.

		Difference	StdErr	P	CI lower	CI upper
P	1-2	29.641	1.236	0.00000	26.209	33.070
Hb	1-3	30.189	1.333	0.00000	26.487	33.890
	1-4	30.543	1.235	0.00000	27.112	33.970
	2-3	0.548	0.400	0.53024	-0.563	1.660
	2-4	0.902	0.345	0.06930	-0.055	1.860
	3-4	0.354	0.198	0.30460	-0.195	0.900
A	1-2	-0.033	0.029	0.67424	-0.113	0.050
Hb	1-3	0.051	0.024	0.18413	-0.016	0.120
	1-4	0.055	0.032	0.33544	-0.034	0.140
	2-3	0.083	0.027	0.02403	0.009	0.160
	2-4	0.088	0.029	0.02850	0.008	0.170
	3-4	0.004	0.025	0.99820	-0.066	0.070

TABLE B.8: Statistical analysis on the explained variance for Experiment 3 using *Hb*. Post-hoc multiple comparisons analysis by eigen-network. Turkey-Kramer correction was used.

		Difference	StdErr	P	CI lower	CI upper
P	1: TPs-Sil	8.123	2.571	0.01213	1.665	14.580
Hb	1: TPs-Rnd	3.382	3.102	0.52988	-4.410	11.170
	1: Rnd-Sil	4.741	3.001	0.27522	-2.799	12.280
	2: TPs-Sil	0.225	0.640	0.93446	-1.383	1.830
	2: TPs-Rnd	0.667	0.538	0.44243	-0.683	2.020
	2: Rnd-Sil	-0.442	0.539	0.69467	-1.797	0.910
	3: TPs-Sil	-0.698	0.344	0.12822	-1.561	0.170
	3: TPs-Rnd	-0.704	0.420	0.23702	-1.760	0.350
	3: Rnd-Sil	0.006	0.353	0.99983	-0.882	0.890
	4: TPs-Sil	-0.092	0.399	0.97102	-1.096	0.910
	4: TPs-Rnd	-0.022	0.464	0.99873	-1.189	1.140
	4: Rnd-Sil	-0.070	0.431	0.98560	-1.154	1.010
A	1: TPs-Sil	0.032	0.045	0.75433	-0.080	0.140
Hb	1: TPs-Rnd	-0.025	0.061	0.90950	-0.179	0.130
	1: Rnd-Sil	0.057	0.047	0.45294	-0.061	0.180
	2: TPs-Sil	0.081	0.044	0.18330	-0.030	0.190
	2: TPs-Rnd	0.123	0.039	0.01281	0.024	0.220
	2: Rnd-Sil	-0.042	0.038	0.53025	-0.138	0.050
	3: TPs-Sil	0.005	0.050	0.99473	-0.120	0.130
	3: TPs-Rnd	0.040	0.053	0.72986	-0.093	0.170
	3: Rnd-Sil	-0.036	0.042	0.68422	-0.142	0.070
	4: TPs-Sil	0.048	0.044	0.54225	-0.064	0.160
	4: TPs-Rnd	0.062	0.048	0.41543	-0.059	0.180
	4: Rnd-Sil	-0.014	0.034	0.90953	-0.100	0.070

TABLE B.9: Statistical analysis on the explained variance for Experiment 3 using *Hb*. Post-hoc multiple comparisons analysis by condition. Turkey-Kramer correction was used.

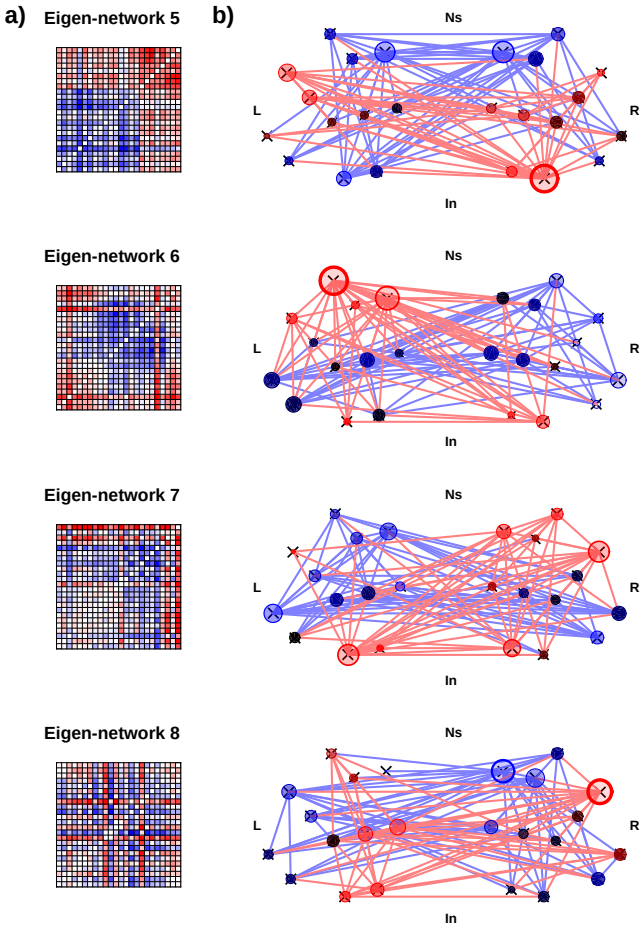


FIGURE B.11: Eigen-networks 5 to 8 for Experiment 3 using *Hb*. The representation uses the same code than in previous figures.

B.3 Appendix to the connectivity analysis for Experiment 2

The eigen-networks 5 to 8 are represented bellow.

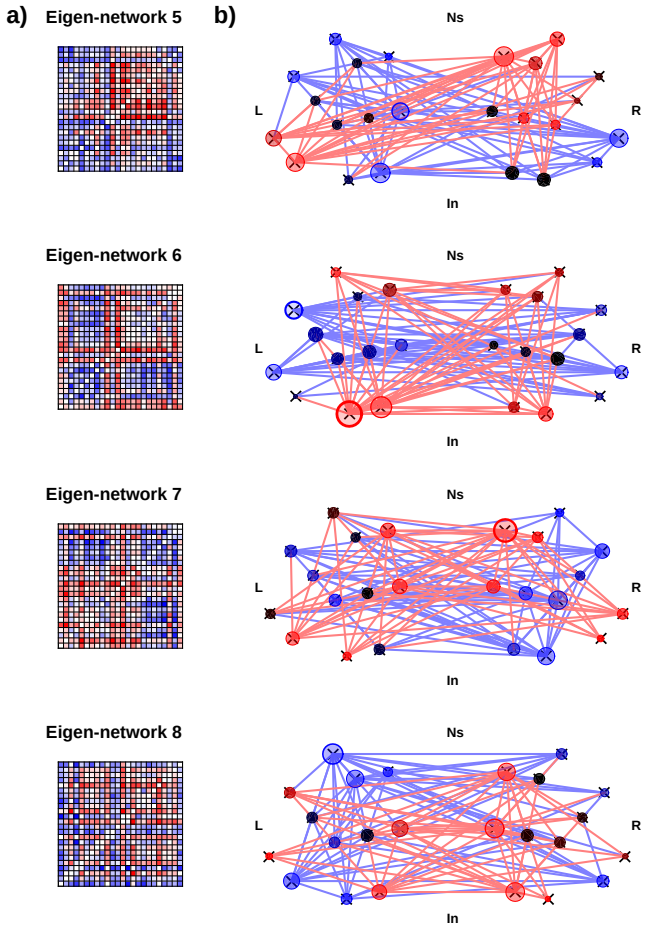
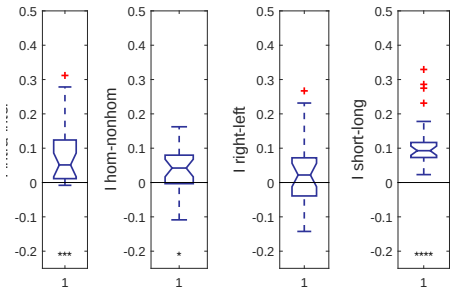
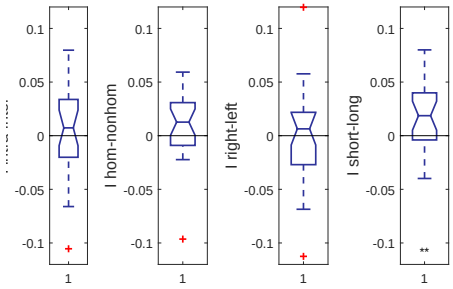


FIGURE B.12: Eigen-networks 5 to 8 for Experiment 2 using HbO_2. The representation uses the same code than in previous figures.

Strength and variability of the connections. Hb



(A) Strength

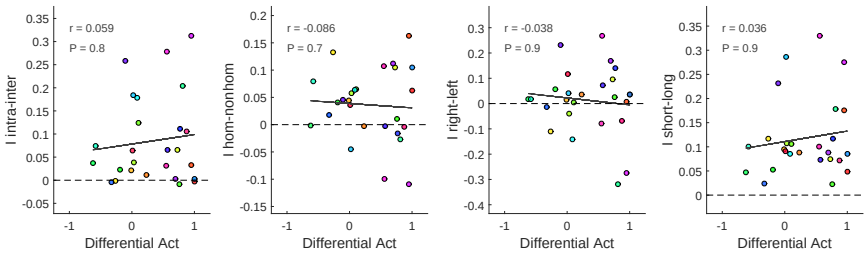


(B) Variability

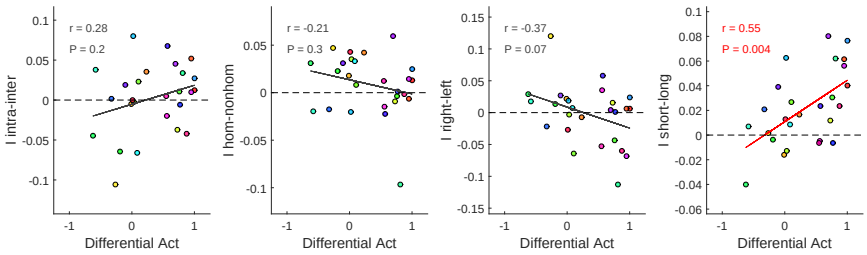
FIGURE B.13: Results for the four indexes defined in terms of the strength **(A)** and temporal variability **(B)** for Experiment 2. Connectivity was estimated using HbO_2. Asterisks bellow represent Bonferroni corrected P-values of the t-tests against zero.

		Mean	SD	CI lower	CI upper	DF	t	P	P_{corr}
Strength	I intra-inter	0.0850	0.0947	0.0468	0.1233	25	4.5792	0.0001	0.0000
Hb	I hom-nonhom	0.0361	0.0663	0.0093	0.0629	25	2.7773	0.0102	0.0410
	I right-left	0.0131	0.1310	-0.0398	0.0661	25	0.5110	0.6138	2.4550
	I short-long	0.1180	0.0803	0.0856	0.1505	25	7.4919	0.0000	0.0000
Variability	I intra-inter	0.0026	0.0431	-0.0148	0.0201	25	0.3103	0.7589	3.0360
Hb	I hom-nonhom	0.0087	0.0316	-0.0041	0.0214	25	1.3961	0.1750	0.7000
	I right-left	-0.0022	0.0463	-0.0209	0.0165	25	-0.2430	0.8100	3.2400
	I short-long	0.0219	0.0307	0.0095	0.0343	25	3.6450	0.0012	0.0050

TABLE B.10: Statistical analysis on the different indexes For Experiment 2 using Hb. T-tests against chance. Bonferroni correction was used.



(A) Strength



(B) Variability

FIGURE B.14: Spearman correlations between task performance and functional connectivity for Experiment 2 using Hb. The differential activation for part-words and words during test blocks (x-axis) is plotted against the different indexes estimated during the familiarization phase (y-axis): $I_{inter-intra}$, $I_{hom-nonhom}$, $I_{right-left}$, $I_{short-long}$. Each dot represent a subject.

.

Principal Components Analysis. Hb

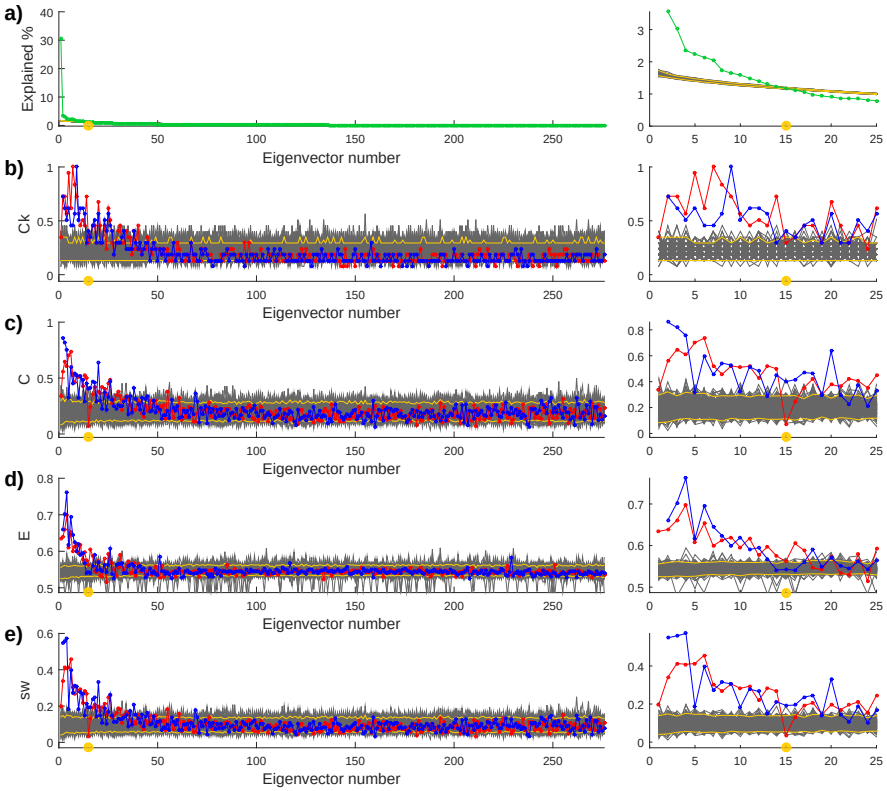


FIGURE B.15: Phase randomization simulations for Experiment 2 using Hb. The right panels show a zoom of the left panels. **a)** Explained variance as a function of the eigenvector number for the real data (green) and the phase randomization simulations (grey). The yellow lines represent the 5% and 95% confidence interval. The yellow dot indicates from which component results can be attributable to noise based on the phase randomization simulations with a significant level of 5%. **b-e)** Graph measures for the positive (red) and negative (blue) parts of the eigen-networks from the real and phase randomized data (grey). **b)** Betweeness centrality. **c)** Cluster coefficient. **d)** Efficiency. **e)** Small-world.

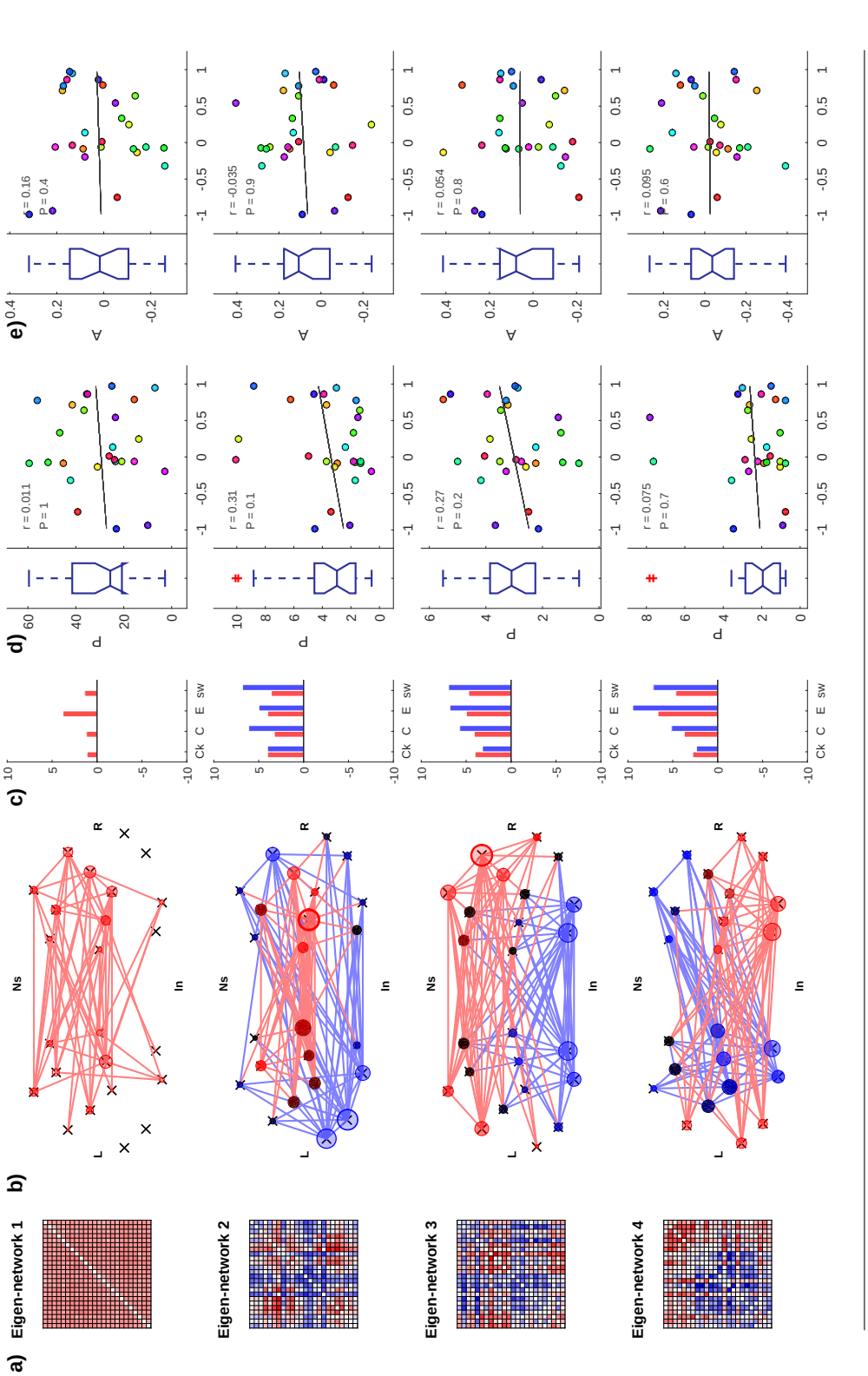


FIGURE B.16: PCA analysis for Experiment 2 using Hb. The first four eigen-networks are shown. **a)** Matrices with the weights to obtain the eigenvectors from the original variables. Positive weights are represent in red and negative weights in blue. **b)** Eigen-networks obtained by thresholding (top 20 %) and binarizing the weights matrices. Crosses represent channels (nodes) in their real relative location, with frontal channels on the top. On each channel a dot is plotted. Its size represents the degree of the node, its colour intensity the cluster coefficient of the node, and the with of its line the centrality of the node. **c)** Z-scored graph measures for the positive and negative parts of the eigen-networks. Betweenness centrality (C_k) . Mean cluster coefficient (C). Global efficiency (E). Small-world (sw). **d)** Spearman correlations between the explained variability (P) for each eigen-network and the differential activation for Part-words and Words. **e)** Spearman correlations between the asymmetry (A), for each eigen-network and the differential activation for Part-words and Words.

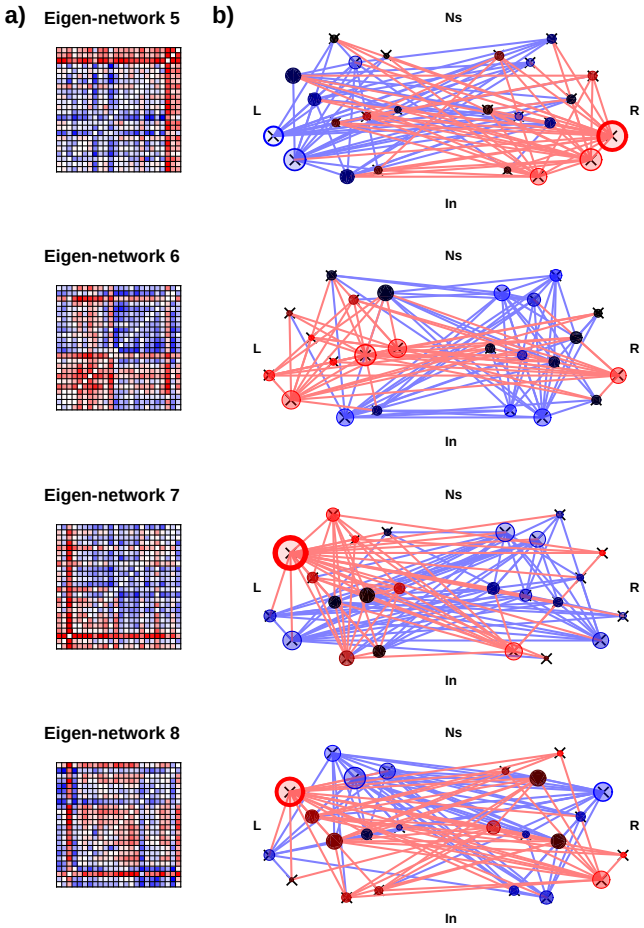


FIGURE B.17: Eigen-networks 5 to 8 for Experiment 2 using *Hb*. The representation uses the same code than in previous figures

Appendix C

Appendix to Chapter 4

C.1 Appendix to Experiment 4

The results using Hb for both the cluster based permutation analysis and mean activation analysis are presented below.

Cluster Based Permutation Analysis.

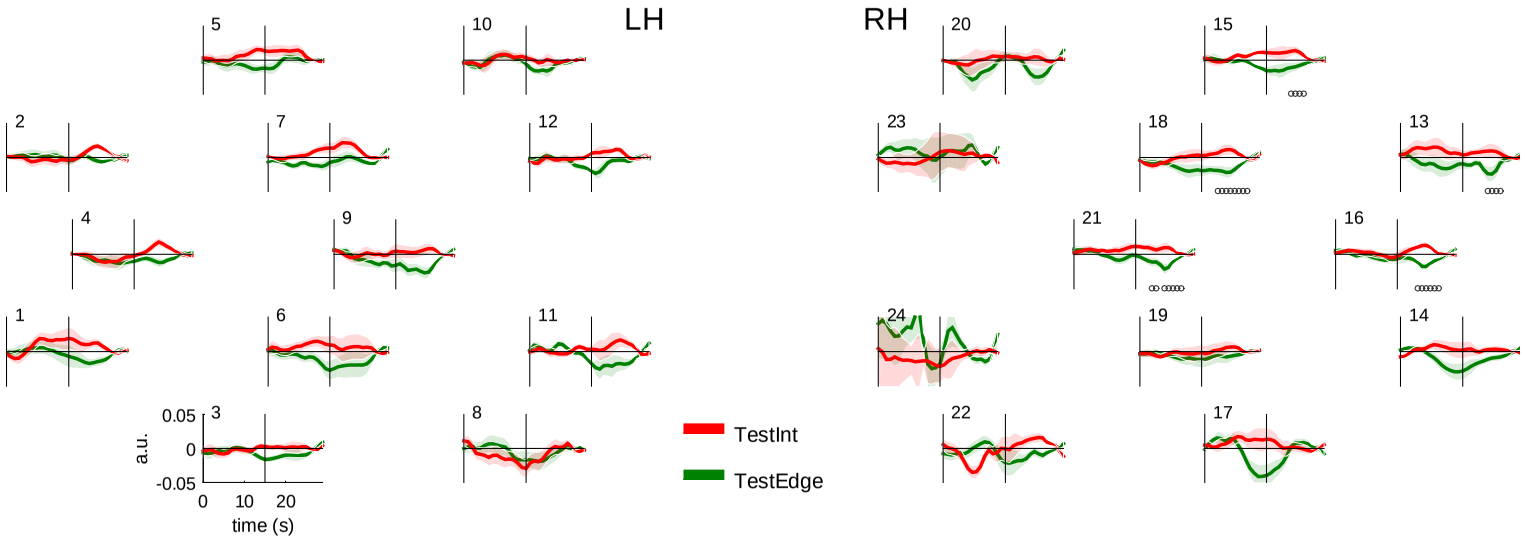


FIGURE C.1: Cluster based permutation analysis for Experiment 4 using *Hb*. HRFs for the Edge switch condition (green) and Internal switch condition (red) during test blocks. Vertical lines are the onset and offset of the stimulus. Marks below show the channels and time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

Mean Activation Analysis.

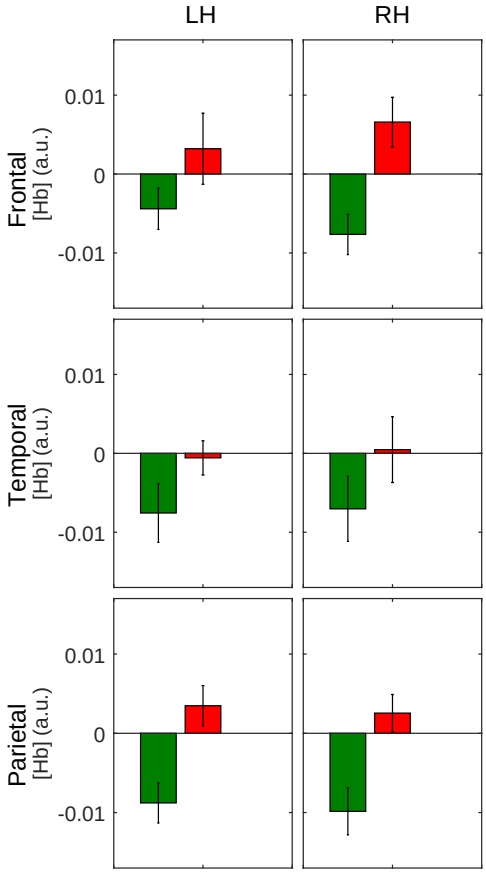


FIGURE C.2: Mean activation analysis for Experiment 4 using Hb. In green activation during the Edge switch condition, and in red activation during Internal switch condition. Error bars represent standard errors

Hb	SumSq	DF	MeanSq	F	P
Condition	0.00273	1	0.00273	19.52137	0.00050
Error (Condition)	0.00210	15	0.00014		
Hemisphere	0.00000	1	0.00000	0.00738	0.93270
Condition: Hemisphere	0.00004	1	0.00004	0.52695	0.47910
Error (Hemisphere)	0.00126	15	0.00008		
Region	0.00030	2	0.00015	1.92636	0.16330
Condition: Region	0.00005	2	0.00003	0.33241	0.71980
Error (Region)	0.00231	30	0.00008		
Hemisphere: Region	0.00001	2	0.00001	0.08200	0.92150
Condition: Hemisphere: Region	0.00004	2	0.00002	0.23374	0.79300
Error (Hemisphere: Region)	0.00263	30	0.00009		

TABLE C.1: Statistical analysis for Experiment 4. 3-ways ANOVA on the mean $[Hb]$ as dependent variables. The hemisphere (left/ right) and the region (frontal/ temporal/ parietal) are within subject factors, and the condition (Edge/ Internal) is the between subjects factor

C.2 Appendix to Experiment 5

The results using Hb for both the cluster based permutation analysis and mean activation analysis are presented below.

Cluster Based Permutation Analysis.

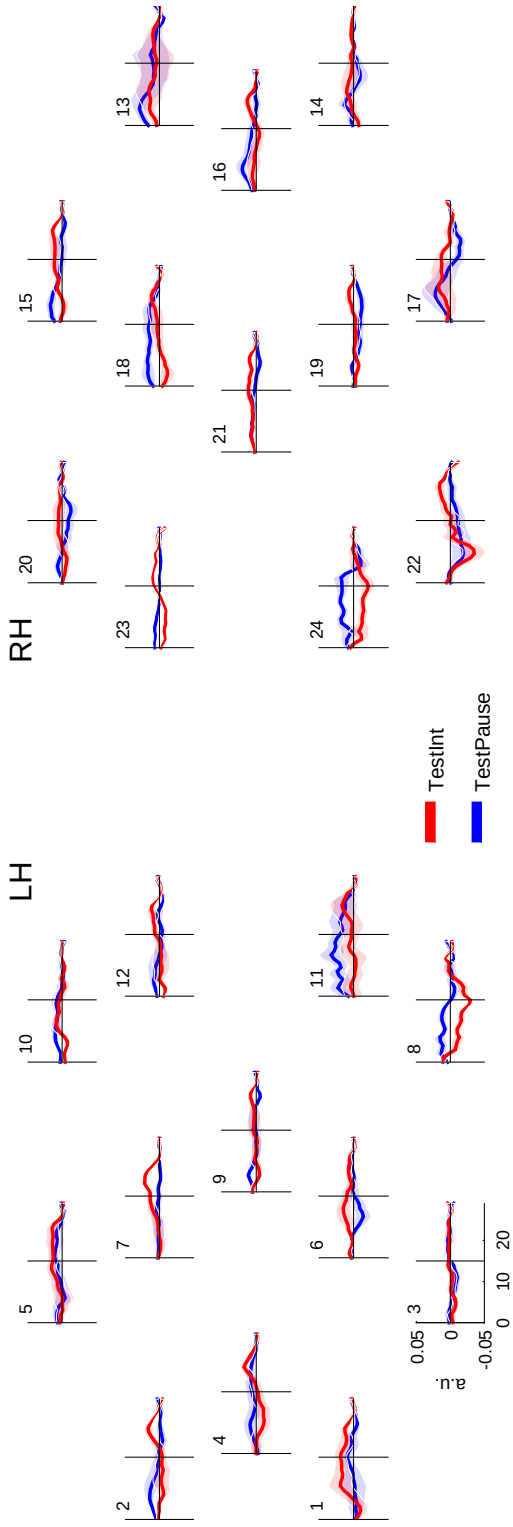


FIGURE C.3: Cluster based permutation analysis for Experiment 5 using *Hb*. HRFs for the Pause switch condition (blue) and Internal switch condition (red) during test blocks. Marks below show the channels and time points in which the cluster based permutation analysis revealed significant differences between conditions. The shaded area represents standard errors.

Mean Activation Analysis.

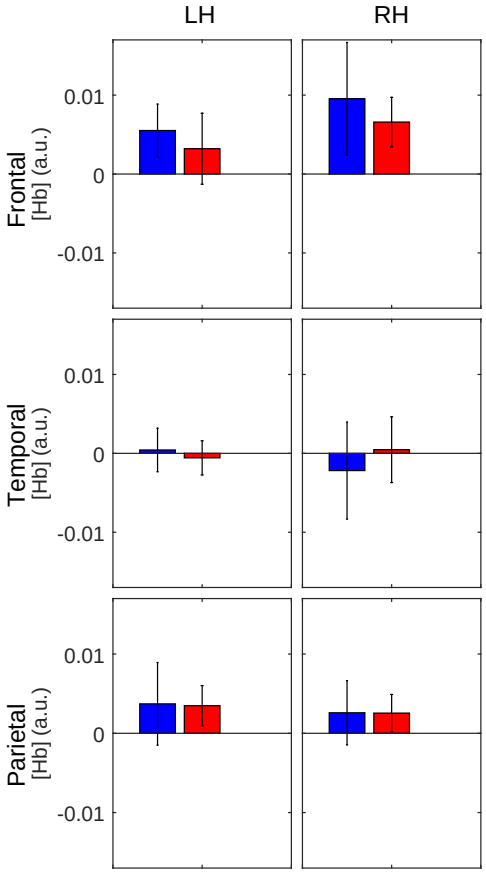


FIGURE C.4: Mean activation analysis for Experiment 4 using *Hb*. In blue activation during Pause switch condition, and in red activation during Internal switch condition. Error bars represent standard errors.

Hb	SumSq	DF	MeanSq	F	P
Condition	0.00010	1	0.00010	0.43536	0.51940
Error	0.00347	15	0.00023		
Hemisphere	0.00006	1	0.00006	0.20589	0.65650
Condition: Hemisphere	0.00000	1	0.00000	0.01068	0.91910
Error (Hemisphere)	0.00419	15	0.00028		
Region	0.00091	2	0.00046	2.64260	0.08770
Condition: Region	0.00010	2	0.00005	0.29019	0.75020
Error (Region)	0.00519	30	0.00017		
Hemisphere: Region	0.00011	2	0.00005	0.74570	0.48300
Condition: Hemisphere: Region	0.00003	2	0.00002	0.22265	0.80170
Error (Hemisphere: Region)	0.00221	30	0.00007		

TABLE C.2: Statistical analysis for Experiment 5 using Hb. 3-ways ANOVA on the mean $[HbO_2]$ and $[Hb]$ as dependent variables. The hemisphere (left/ right) and the region (frontal/ temporal/ parietal) are within subject factors, and the condition (Edge/ Internal) is the between subjects factor

Bibliography

Abla, Dilshat, Kentaro Katahira, and Kazuo Okanoya (2008). "On-line assessment of statistical learning by event-related potentials". In: *Journal of Cognitive Neuroscience* 20.6, pp. 952–64. ISSN: 0898-929X. DOI: 10.1162/jocn.2008.20058.

Abla, Dilshat and Kazuo Okanoya (2009). "Visual statistical learning of shape sequences: An ERP study". In: *Neuroscience Research* 64.2, pp. 185–190. ISSN: 01680102. DOI: 10.1016/j.neures.2009.02.013.

Adachi, Yusuke et al. (2012). "Functional connectivity between anatomically unconnected areas is shaped by collective network-level effects in the macaque cortex". In: *Cerebral Cortex* 22.7, pp. 1586–1592. ISSN: 10473211. DOI: 10.1093/cercor/bhr234.

Allen, Elena A. et al. (2014). "Tracking whole-brain connectivity dynamics in the resting state". In: *Cerebral Cortex* 24.3, pp. 663–676. ISSN: 10473211. DOI: 10.1093/cercor/bhs352.

Aslin, R. N. and E. L. Newport (2012). "Statistical Learning: From Acquiring Specific Items to Forming General Rules". In: *Current Directions in Psychological Science* 21.3, pp. 170–176. ISSN: 0963-7214. DOI: 10.1177/0963721412436806. arXiv: NIHMS150003.

Aslin, R. N., J. R. Saffran, and E. L. Newport (1998). "Computation of Conditional Probability Statistics by 8-Month-Old Infants". In: *Psychological Science* 9.4, pp. 321–324. ISSN: 0956-7976. DOI: 10.1111/1467-9280.00063.

Assaf, Michal et al. (2010). "Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients". In: *NeuroImage* 53.1, pp. 247–256. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.05.067.

Baddeley, Alan and Graham Hitch (1974). "Working memory." In: *Psychology of Learning and Motivation* 8, pp. 47–89. ISSN: 1879-0445. DOI: 10.1016/j.cub.2009.12.014.

Bahlmann, Jörg, Ricarda I. Schubotz, and Angela D. Friederici (2008). "Hierarchical artificial grammar processing engages Broca's area". In: *NeuroImage* 42.2, pp. 525–534. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2008.04.249.

Barrett, H C, W E Frankenhuis, and A Wilke (2008). "Adaptation to moving targets: Culture/gene coevolution, not either/or". In: *Behavioral and Brain Sciences* 31.5, pp. 511–+. ISSN: 0140-525X. DOI: 10.1017/S0140525x08005013.

Barttfeld, Pablo et al. (2014). "Signature of consciousness in the dynamics of resting-state brain activity". In: *Proceedings of the National Academy of Sciences* 112.3, p. 201418031. ISSN: 0027-8424. DOI: 10.1073/pnas.1418031112.

Bassett, Danielle S. and Michael S. Gazzaniga (2011). "Understanding complexity in the human brain". In: *Trends in Cognitive Sciences* 15.5, pp. 200–209. ISSN: 13646613. DOI: 10.1016/j.tics.2011.03.006.

Bassett, Danielle S et al. (2011). "Dynamic reconfiguration of human brain networks during learning." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.18, pp. 7641–6. ISSN: 1091-6490. DOI: 10.1073/pnas.1018985108. arXiv: 1010.3775.

Bassett, Danielle S. et al. (2015). "Learning-Induced Autonomy of Sensorimotor Systems". In: *Nature neuroscience* 18.5, pp. 744–751. ISSN: 1097-6256. DOI: 10.1038/nn.3993. arXiv: arXiv:1403.6034v1.

Benavides-Varela, S. and J. Mehler (2012). "No Title". PhD thesis. SISSA, pp. 47–65.

Benavides-Varela, S. et al. (2012). "Newborn's brain activity signals the origin of word memories". In: *Proceedings of the National Academy of Sciences* 109.44, pp. 17908–17913. ISSN: 0027-8424. DOI: 10.1073/pnas.1205413109.

Benavides-Varela, Silvia and Jacques Mehler (2015). "Verbal Positional Memory in 7-Month-Olds". In: *Child Development* 86.1, pp. 209–223. ISSN: 14678624. DOI: 10.1111/cdev.12291.

Benavides-Varela, Silvia et al. (2011). "Memory in the Neonate Brain". In: *PLoS ONE* 6.11. Ed. by Christine Jasoni, e27497. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0027497.

Berwick, Robert C. et al. (2013). *Evolution, brain, and the nature of language*. DOI: 10.1016/j.tics.2012.12.002.

Bickerton, Derek (1992). "Language & species." In: *Chicago*, pp. 1–12, 255–257. ISBN: 0226046109.

Boveroux, Pierre, Audrey Vanhaudenhuyse, and Christophe Phillips (2010). "Breakdown of within- and between-network Resting State during Propofol-induced Loss of Consciousness". In: *Anesthesiology* 113.5, pp. 1038–1053.

Brainard, David H. (1997). "The Psychophysics Toolbox". In: *Spatial Vision* 10, pp. 433–436. ISSN: 0169-1015. DOI: 10.1163/156856897X00357. arXiv: arXiv:1011.1669v3.

Braun, Urs et al. (2015). "Dynamic reconfiguration of frontal brain networks during executive cognition in humans." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.37, pp. 11678–83. ISSN: 1091-6490. DOI: 10.1073/pnas.1422487112. arXiv: arXiv:1408.1149.

Bryant, Gregory A. and H. Clark Barrett (2007). "Recognizing intentions in infant-directed speech: Evidence for universals". In: *Psychological Science* 18.8, pp. 746–751. ISSN: 09567976. DOI: 10.1111/j.1467-9280.2007.01970.x.

Buiatti, Marco, Marcela Peña, and Ghislaine Dehaene-Lambertz (2009). "Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses". In: *NeuroImage* 44.2, pp. 509–519. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2008.09.015.

Bulf, Hermann, Scott P. Johnson, and Eloisa Valenza (2011). "Visual statistical learning in the newborn infant". In: *Cognition* 121.1, pp. 127–132. ISSN: 00100277. DOI: 10.1016/j.cognition.2011.06.010.

Bullmore, Ed and Olaf Sporns (2012). "The economy of brain network organization". In: *Nature Reviews Neuroscience* 13.MAY, pp. 336–349. ISSN: 1471-003X. DOI: 10.1038/nrn3214.

Bullmore, Edward T., Olaf Sporns, and Sara A Solla (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems." In: *Nature reviews. Neuroscience* 10.3, pp. 186–98. ISSN: 1471-0048. DOI: 10.1038/nrn2575.

Celsis, P et al. (1999). "Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones." In: *NeuroImage* 9.1, pp. 135–144. ISSN: 1053-8119. DOI: 10.1006/nimg.1998.0389.

Chang, Catie et al. (2013). "Association between heart rate variability and fluctuations in resting-state functional connectivity". In: *NeuroImage* 68, pp. 93–104. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.11.038. arXiv: NIHMS150003.

Chomsky, Noam (1957). *Syntactic Structures*. Vol. 7. 2, p. 116. ISBN: 3110172798. DOI: 10.1111/j.1467-9612.2004.00004.x.

— (1995). "The Minimalist Program". In: *Journal of Linguistics* 34.1, pp. 213–226. ISSN: 00222267. DOI: 10.1017/S0022226797006889.

Christiansen, M H and N Chater (2008). "Language as shaped by the brain". In: *Behavioral and Brain Sciences* 31.5, pp. 458–489. ISSN: 0140-525X. DOI: 10.1017/S0140525X08004998.

Christiansen, Morten H, Christopher M Conway, and Luca Onnis (2012). "Similar Neural Correlates for Language and Sequential Learning: Evidence from Event-Related Brain Potentials." In: *Language and cognitive processes* 27.2, pp. 231–256. ISSN: 0169-0965. DOI: 10.1080/01690965.2011.606666.

Christiansen, Morten H. et al. (2006). "Statistical learning within and between modalities". In: *Psychological Science* 17.10, pp. 905–912. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2006.01801.x.

Christophe, Anne et al. (1994). "Do infants percieve word boundaries? An empirical study of the bootstraőőing of lexical acquisition". In: *Journal of Acoustical Society of America* 95.May, pp. 1570–1580. ISSN: 00014966. DOI: 10.1121/1.408544.

Cohen, Jonathan et al. (1993). "PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers". In: *Behavior Research Methods, Instruments, & Computers* 25.2, pp. 257–271. ISSN: 0743-3808. DOI: 10.3758/BF03204507.

Cole, M W et al. (2013). "Multi-task connectivity reveals flexible hubs for adaptive task control". In: *Nature Neuroscience* 16.9, pp. 1348–1355. ISSN: 1097-6256. DOI: 10.1038/nn.3470.

Cole, Michael W. et al. (2014). "Intrinsic and task-evoked network architectures of the human brain". In: *Neuron* 83.1, pp. 238–251. ISSN: 10974199. DOI: 10.1016/j.neuron.2014.05.014. arXiv: NIHMS150003.

Conway, Christopher M and Morten H Christiansen (2001). "Review: Sequential learning in non-human primates". In: *Trends in Cognitive Sciences* 5.12, pp. 539–546. ISSN: 1364-6613. DOI: 10.1016/S1364-6613(00)01800-3.

Cooper, Robert J. et al. (2012). "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy". In: *Frontiers in Neuroscience* 6.OCT, pp. 1–10. ISSN: 16624548. DOI: 10.3389/fnins.2012.00147.

Cope, Mark and David T Delpy (1988). "System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination". In: *Medical & Biological Engineering & Computing* 26.3, pp. 289–294. ISSN: 0140-0118. DOI: 10.1007/BF02447083.

Cunillera, Toni et al. (2009). "Time course and functional neuroanatomy of speech segmentation in adults". In: *NeuroImage* 48.3, pp. 541–553. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.06.069.

Davison, Elizabeth N. et al. (2015). "Brain Network Adaptability across Task States". In: *PLoS Computational Biology* 11.1. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1004029. arXiv: 1407.8234.

DeCasper, a J and W P Fifer (1980). "Of human bonding: newborns prefer their mothers' voices." In: *Science (New York, N.Y.)* 208.4448, pp. 1174–1176. ISSN: 0036-8075. DOI: 10.1126/science.7375928.

DeCasper, Anthony J. and Melanie J. Spence (1986). "Prenatal maternal speech influences newborns' perception of speech sounds". In: *Infant Behavior and Development* 9.2, pp. 133–150. ISSN: 01636383. DOI: 10.1016/0163-6383(86)90025-1.

Deco, Gustavo, Viktor K Jirsa, and Anthony R McIntosh (2011). "Emerging concepts for the dynamical organization of resting-state activity in the brain." In: *Nature reviews. Neuroscience* 12.1, pp. 43–56. ISSN: 1471-003X. DOI: 10.1038/nrn2961.

Dehaene, Stanislas et al. (2015). "The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees". In: *Neuron* 88.1, pp. 2–19. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.09.019.

Dehaene-Lambertz, G and M Pena (2001). "Electrophysiological evidence for automatic phonetic processing in neonates." In: *Neuroreport* 12.14, pp. 3155–3158. ISSN: 0959-4965. DOI: 10.1097/00001756-200110080-00034.

Dehaene-Lambertz, G. and E. S. Spelke (2015). "The Infancy of the Human Brain". In: *Neuron* 88.1, pp. 93–109. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.09.026.

Dehaene-Lambertz, Ghislaine, Stanislas Dehaene, and Lucie Hertz-Pannier (2002). "Functional Neuroimaging of Speech Perception in Infants". In: *Science* 298.5600, pp. 2013–2015. ISSN: 1095-9203. DOI: 10.1126/science.1077066.

Dehaene-Lambertz, Ghislaine, Lucie Hertz-Pannier, and Jessica Dubois (2006). "Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants". In: *Trends in Neurosciences* 29.7, pp. 367–373. ISSN: 01662236. DOI: 10.1016/j.tins.2006.05.011.

Dehaene-Lambertz, Ghislaine et al. (2006a). "Functional organization of perisylvian activation during presentation of sentences in preverbal infants". In: *Proceedings of the National Academy of Sciences* 103.38, pp. 14240–14245. ISSN: 0027-8424. DOI: 10.1073/pnas.0606302103.

Dehaene-Lambertz, Ghislaine et al. (2006b). "Functional segregation of cortical language areas by sentence repetition". In: *Human Brain Mapping* 27.5, pp. 360–371. ISSN: 10659471. DOI: 10.1002/hbm.20250.

Doria, Valentina et al. (2010). "Emergence of resting state networks in the preterm human brain". In: *Proceedings of the National Academy of Sciences* 107.46, pp. 20015–20020. ISSN: 0027-8424. DOI: 10.1073/pnas.1007921107.

Dubois, J. et al. (2009). "Structural asymmetries in the infant language and sensorimotor networks". In: *Cerebral Cortex* 19.2, pp. 414–423. ISSN: 10473211. DOI: 10.1093/cercor/bhn097.

Dubois, J. et al. (2014). "The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants". In: *Neuroscience* 276, pp. 48–71. ISSN: 18737544. DOI: 10.1016/j.neuroscience.2013.12.044.

Dubois, Jessica et al. (2008). "Asynchrony of the early maturation of white matter bundles in healthy infants: Quantitative landmarks revealed noninvasively by diffusion tensor imaging". In: *Human Brain Mapping* 29.1, pp. 14–27. ISSN: 10659471. DOI: 10.1002/hbm.20363.

Durand, Frédo and Julie Dorsey (2002). "Fast bilateral filtering for the display of high-dynamic-range images". In: *ACM Transactions on Graphics* 21.3, pp. 257–266. ISSN: 07300301. DOI: 10.1145/566654.566574.

Dutoit, T. et al. (1996). "The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes". In: *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*. Vol. 3. IEEE, pp. 1393–1396. ISBN: 0-7803-3555-4. DOI: 10.1109/ICSLP.1996.607874.

Ebbinghaus, Hermann (1885). "Memory: A Contribution to Experimental Psychology". In: *Memory: A Contribution to Experimental Psychology* 15, pp. 1–7. ISSN: 0009-4978. DOI: 52,281-302.

Edwards, Laura A. et al. (2015). "Functional brain organization for number processing in pre-verbal infants". In: *Developmental Science*, pp. 1–13. ISSN: 14677687. DOI: 10.1111/desc.12333.

Eimas, P D and J L Miller (1980). "Contextual effects in infant speech perception". In: *Science* 209.4461, pp. 1140–1141. ISSN: 00368075. DOI: 10.1126/science.7403875.

Eimas, PD et al. (1971). "Speech perception in infants". In: *Science* 171.3968, pp. 303–306.

Emberson, Lauren L., John E. Richards, and Richard N. Aslin (2015). "Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months". In: *Proceedings of the National Academy of Sciences* 112.31, pp. 9585–9590. ISSN: 0027-8424. DOI: 10.1073/pnas.1510343112.

Endress, Ansgar D. and Marc D. Hauser (2010). "Word segmentation with universal prosodic cues". In: *Cognitive Psychology* 61.2, pp. 177–199. ISSN: 00100285. DOI: 10.1016/j.cogpsych.2010.05.001.

Endress, Ansgar D and Jacques Mehler (2009a). "Primitive computations in speech processing." In: *Quarterly journal of experimental psychology (2006)* 62.11, pp. 2187–209. ISSN: 1747-0226. DOI: 10.1080/17470210902783646.

Endress, Ansgar D. and Jacques Mehler (2009b). "The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words". In: *Journal of Memory and Language* 60.3, pp. 351–367. ISSN: 0749596X. DOI: 10.1016/j.jml.2008.10.003.

Endress, Ansgar D, Marina Nespor, and Jacques Mehler (2009). "Perceptual and memory constraints on language acquisition." In: *Trends in cognitive sciences* 13.8, pp. 348–53. ISSN: 1364-6613. DOI: 10.1016/j.tics.2009.05.005.

Endress, Ansgar D, Brian J Scholl, and Jacques Mehler (2005). "The role of salience in the extraction of algebraic rules." In: *Journal of experimental psychology. General* 134.3, pp. 406–19. ISSN: 0096-3445. DOI: 10.1037/0096-3445.134.3.406.

Fekete, Tomer et al. (2011). "The NIRS Analysis Package: Noise Reduction and Statistical Inference". In: *PLoS ONE* 6.9. Ed. by Xi-Nian Zuo, e24322. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0024322.

Fernald, Anne (1989). "Intonation and Communicateive Intent in Mothers' Speech to Infants: Is the Melody the Message?" In: *Child development* 60.6, pp. 1497–1510. ISSN: 0009-3920. DOI: 10.1111/1467-8624.ep9772504.

— (1992). "Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective". In: *The Adapted Mind*, pp. 391–428. ISSN: 0169-3867. DOI: 10.1007/BF00852474.

Ferrari, Marco and Valentina Quaresima (2012). "A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application". In: *NeuroImage* 63.2, pp. 921–935. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.03.049.

Ferry, Alissa L., Susan J. Hespos, and Sandra R. Waxman (2010). "Categorization in 3- and 4-month-old infants: An advantage of words over tones". In: *Child Development* 81.2, pp. 472–479. ISSN: 00093920. DOI: 10.1111/j.1467-8624.2009.01408.x. arXiv: NIHMS150003.

Ferry, Alissa L, Susan J Hespos, and Sandra R Waxman (2013). "Nonhuman primate vocalizations support categorization in very young human infants." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.38, pp. 15231–5. ISSN: 1091-6490. DOI: 10.1073/pnas.1221166110.

Ferry, Alissa L. et al. (2016). "On the edge of language acquisition: inherent constraints on encoding multisyllabic sequences in the neonate brain". In: *Developmental Science* 19.3, pp. 488–503. ISSN: 1363755X. DOI: 10.1111/desc.12323.

Fiser, József and Richard N Aslin (2002). "Statistical learning of new visual feature combinations by infants." In: *Proceedings of the National Academy of Sciences of the United States of America* 99.24, pp. 15822–6. ISSN: 0027-8424. DOI: 10.1073/pnas.232472899.

Fitch, W T (2008). "Co-evolution of phylogeny and glossogeny: There is no "logical problem of language evolution"". In: *Behavioral and Brain Sciences* 31.5, pp. 521–522. ISSN: 0140525X (ISSN). DOI: 10.1017/S0140525X08005128.

Floel, Agnes et al. (2009). "White matter integrity in the vicinity of Broca's area predicts grammar learning success". In: *NeuroImage* 47.4, pp. 1974–1981. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.05.046.

Fodor, Jerry A. and Zenon W. Pylyshyn (1988). "Connectionism and cognitive architecture: A critical analysis". In: *Cognition* 28.1-2, pp. 3–71. ISSN: 00100277. DOI: 10.1016/0010-0277(88)90031-5. arXiv: arXiv:1011.1669v3.

Fournier, Lisa R et al. (2014). "On the importance of being first: serial order effects in the interaction between action plans and ongoing actions." In: *Psychonomic bulletin & review* 21.1, pp. 163–9. ISSN: 1531-5320. DOI: 10.3758/s13423-013-0486-0.

Frank, Stefan L, Rens Bod, and Morten H Christiansen (2012). "How hierarchical is language use?" In: *Proceedings of the Royal Society B: Biological Sciences* 279.1747, pp. 4522–4531. ISSN: 0962-8452. DOI: 10.1098/rspb.2012.1741.

Fransson, Peter et al. (2007). "Resting-state networks in the infant brain." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.39, pp. 15531–6. ISSN: 0027-8424. DOI: 10.1073/pnas.0704380104.

Fransson, Peter et al. (2011). "The functional architecture of the infant brain as revealed by resting-state fMRI". In: *Cerebral Cortex* 21.1, pp. 145–154. ISSN: 10473211. DOI: 10.1093/cercor/bhq071.

Friederici, Angela D. and Kai Alter (2004). *Lateralization of auditory language functions: A dynamic dual pathway model*. DOI: 10.1016/S0093-934X(03)00351-1.

Friederici, Angela D., Manuela Friedrich, and Anne Christophe (2007). "Brain Responses in 4-Month-Old Infants Are Already Language Specific". In: *Current Biology* 17.14, pp. 1208–1211. ISSN: 09609822. DOI: 10.1016/j.cub.2007.06.011.

Friederici, Angela D. and Sarah M E Gierhan (2013). *The language network*. DOI: 10.1016/j.conb.2012.10.002.

Friederici, Angela D, Regine Oberecker, and Jens Brauer (2012). "Neurophysiological preconditions of syntax acquisition." In: *Psychological research* 76.2, pp. 204–11. ISSN: 1430-2772. DOI: 10.1007/s00426-011-0357-0.

Friederici, Angela D et al. (2006). "The brain differentiates human and non-human grammars: functional localization and structural connectivity." In: *Proceedings of the National Academy of Sciences of the United States of America* 103.7, pp. 2458–63. ISSN: 0027-8424. DOI: 10.1073/pnas.0509389103.

Frost, Ram et al. (2015). "Domain generality versus modality specificity: The paradox of statistical learning". In: *Trends in Cognitive Sciences* 19.3, pp. 117–125. ISSN: 1879307X. DOI: 10.1016/j.tics.2014.12.010.

G. F. Marcus, S. Vijayan, S. Bandi Rao, P. M. Vishton (1999). "Rule Learning by Seven-Month-Old Infants". In: *Science* 283.5398, pp. 77–80. ISSN: 00368075. DOI: 10.1126/science.283.5398.77.

Gandour, Jackson et al. (2004). "Hemispheric roles in the perception of speech prosody". In: *NeuroImage* 23.1, pp. 344–357. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2004.06.004.

Gerken, LouAnn (2010). "Infants use rational decision criteria for choosing among models of their input". In: *Cognition* 115.2, pp. 362–366. ISSN: 00100277. DOI: 10.1016/j.cognition.2010.01.006.

Gerken, LouAnn, Peter W. Jusczyk, and Denise R. Mandel (1994). "When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases". In: *Cognition* 51.3, pp. 237–265. ISSN: 00100277. DOI: 10.1016/0010-0277(94)90055-8.

Gervain, Judit and Janet F Werker (2013). "Prosody cues word order in 7-month-old bilingual infants." In: *Nature communications* 4, p. 1490. ISSN: 2041-1723. DOI: 10.1038/ncomms2430.

Gervain, Judit et al. (2008a). "Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study." In: *Cognitive psychology* 57.1, pp. 56–74. ISSN: 1095-5623. DOI: 10.1016/j.cogpsych.2007.12.001.

Gervain, Judit et al. (2008b). "The neonate brain detects speech structure." In: *Proceedings of the National Academy of Sciences of the United States of America* 105.37, pp. 14222–14227. ISSN: 0027-8424. DOI: 10.1073/pnas.0806530105.

Gervain, Judit et al. (2011). "Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium". In: *Developmental Cognitive Neuroscience* 1.1, pp. 22–46. ISSN: 18789293. DOI: 10.1016/j.dcn.2010.07.004. arXiv: arXiv:1011.1669v3.

Glasel, H. et al. (2011). "A robust cerebral asymmetry in the infant brain: The rightward superior temporal sulcus". In: *NeuroImage* 58.3, pp. 716–723. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2011.06.016.

Gredebäck, Gustaf and Annika Melinder (2010). "Infants' understanding of everyday social interactions: A dual process account". In: *Cognition* 114.2, pp. 197–206. ISSN: 00100277. DOI: 10.1016/j.cognition.2009.09.004.

Greenberg, Joseph H. (1957). *Essays in linguistics*. University. Chicago, IL, US, vii 108 pp.

Gulya, M et al. (1998). "Memory Processing of a Serial List by Young Infants". In: *Psychological Science* 9.4, pp. 303–307. ISSN: 0956-7976. DOI: 10.1111/1467-9280.00060.

Gulya, Michelle et al. (2001). "Infants' long-term memory for a serial list: Recognition and reactivation". In: *Developmental Psychobiology* 38.3, pp. 174–185. ISSN: 0012-1630. DOI: 10.1002/dev.1011.

Gupta, Prahlad (2003). "Primacy and recency in nonword repetition." In: *Memory (Hove, England)* 13.3-4, pp. 318–324. ISSN: 0965-8211. DOI: 10.1080/09658210344000350.

Gupta, Prahlad et al. (2005). "Serial position effects in nonword repetition". In: *Journal of Memory and Language* 53.1, pp. 141–162. ISSN: 0749596X. DOI: 10.1016/j.jml.2004.12.002.

Halle, Morris and Jean-Roger Vergnaud (1987). *An essay on stress*. MIT press.

Handwerker, Daniel A. et al. (2012). "Periodic changes in fMRI connectivity". In: *NeuroImage* 63.3, pp. 1712–1719. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2012.06.078. arXiv: NIHMS150003.

Hauser, Marc D, Noam Chomsky, and William T S Fitch (2002). "The faculty of language: what is it, who has it, and how did it evolve?" In: *Science* 298.5598, pp. 1569–79. ISSN: 1095-9203. DOI: 10.1126/science.298.5598.1569.

Hauser, Marc D., Elissa L. Newport, and Richard N. Aslin (2001). "Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins". In: *Cognition* 78.3, pp. 53–64. ISSN: 00100277. DOI: 10.1016/S0010-0277(00)00132-3.

Hauser, Marc D., Daniel Weiss, and Gary Marcus (2002). "RETRACTED: Rule learning by cotton-top tamarins". In: *Cognition* 86.1, B15–B22. ISSN: 00100277. DOI: 10.1016/S0010-0277(02)00139-7.

Hauser, Marc D. et al. (2014). "The mystery of language evolution". In: *Frontiers in Psychology* 5.MAY, pp. 1–12. ISSN: 16641078. DOI: 10.3389/fpsyg.2014.00401.

Hawthorne, Kara and LouAnn Gerken (2014). "From pauses to clauses: Prosody facilitates learning of syntactic constituency". In: *Cognition* 133.2, pp. 420–428. ISSN: 18737838. DOI: 10.1016/j.cognition.2014.07.013.

Hay, Jessica S F and Randy L Diehl (2007). "Perception of rhythmic grouping: testing the iambic/trochaic law." In: *Perception & psychophysics* 69.1, pp. 113–122. ISSN: 0031-5117. DOI: 10.3758/BF03194458.

Hayes, Bruce (1985). "Iambic and trochaic rhythm in stress rules". In: *Annual Meeting of the Berkeley Linguistics Society* 11, pp. 429–446.

— (1989). *The prosodic hierarchy in meter*.

— (1995). *Metrical Stress Theory: Principles and Case Studies*, p. 472. ISBN: 0226321037.

He, Y and A Evans (2010). "Graph theoretical modeling of brain connectivity." In: *Current opinion in neurology* 23.4, pp. 341–350. ISSN: 1350-7540. DOI: 10.1097/WCO.0b013e32833aa567.

Heine, Lizette et al. (2012). "Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states". In: *Frontiers in Psychology* 3.AUG. ISSN: 16641078. DOI: 10.3389/fpsyg.2012.00295.

Henson, R N a (1999). "Coding position in short-term memory". In: *International Journal of Psychology* 34, pp. 403–409. ISSN: 0020-7594. DOI: 10.1080/002075999399756.

Henson, R N a et al. (1996). "Unchained memory : Error patterns rule out chaining models of immediate serial recall". In: *Quarterly Journal of Experimental Psychology* 49.1, pp. 80–115. ISSN: 0272-4987. DOI: 10.1080/713755612.

Henson, Richard N.A. (1998). "Short-Term Memory for Serial Order: The Start-End Model". In: *Cognitive psychology* 36.2, pp. 73–137. ISSN: 0010-0285. DOI: 10.1006/cogp.1998.0685.

Hermundstad, Ann M et al. (2013). "Structural foundations of resting-state and task-based functional connectivity in the human brain." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.15, pp. 6169–74. ISSN: 1091-6490. DOI: 10.1073/pnas.1219562110.

Hess, E H and J M Polt (1964). "Pupil Size in Relation to Mental Activity during Simple Problem-Solving." In: *Science (New York, N.Y.)* 143.3611, pp. 1190–1192. ISSN: 0036-8075. DOI: 10.1126/science.143.3611.1190.

Hess, Eckhard H. and James M. Polt (1960). "Pupil size as related to interest value of visual stimuli". In: *Science* 132, pp. 349–50. ISSN: 0036-8075. DOI: 10.1126/science.132.3423.349.

Hirsh-Pasek, Kathy et al. (1987). "Clauses are perceptual units for young infants". In: *Cognition* 26.3, pp. 269–286. ISSN: 00100277. DOI: 10.1016/S0010-0277(87)80002-1.

Hochmann, Jean R??my (2013). "Word frequency, function words and the second gavagai problem". In: *Cognition* 128.1, pp. 13–25. ISSN: 18737838. DOI: 10.1016/j.cognition.2013.02.014.

Homae, Fumitaka (2014). "A brain of two halves: Insights into interhemispheric organization provided by near-infrared spectroscopy". In: *NeuroImage* 85, pp. 354–362. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.06.023.

Homae, Fumitaka et al. (2006). "The right hemisphere of sleeping infant perceives sentential prosody". In: *Neuroscience Research* 54.4, pp. 276–280. ISSN: 01680102. DOI: 10.1016/j.neures.2005.12.006.

Homae, Fumitaka et al. (2007). "Prosodic processing in the developing brain". In: *Neuroscience Research* 59.1, pp. 29–39. ISSN: 01680102. DOI: 10.1016/j.neures.2007.05.005.

Homae, Fumitaka et al. (2010). "Development of global cortical networks in early infancy." In: *The Journal of Neuroscience* 30.14, pp. 4877–4882. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.5618-09.2010.

Honey, C J et al. (2009). "Predicting human resting-state functional connectivity from structural connectivity." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.6, pp. 2035–40. ISSN: 1091-6490. DOI: 10.1073/pnas.0811168106.

Horovitz, Silvina G et al. (2009). "Decoupling of the brain's default mode network during deep sleep." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.27, pp. 11376–11381. ISSN: 0027-8424. DOI: 10.1073/pnas.0901435106.

Hunter, Michael A; Ames, Elinor W (1988). "A multifactor model of infant preferences for novel and familiar stimuli". In: *Advances in Infancy Research* 5, pp. 69–95. ISSN: 07329598.

Huppert, Theodore J et al. (2009). "HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain". In: *Applied optics* 48.10, pp. 280–298. ISSN: 08966273. DOI: 10.1016/j.drugalcdep.2008.02.002.A. arXiv: NIHMS150003.

Hurlstone, Mark J, Graham J Hitch, and Alan D Baddeley (2014). "Memory for serial order across domains: An overview of the literature and directions for future research." In: *Psychological bulletin* 140.2, pp. 339–73. ISSN: 1939-1455. DOI: 10.1037/a0034221.

Hutchison, R. Matthew et al. (2013a). "Dynamic functional connectivity: Promise, issues, and interpretations". In: *NeuroImage* 80, pp. 360–378. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.05.079. arXiv: NIHMS150003.

Hutchison, R. Matthew et al. (2013b). "Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques". In: *Human Brain Mapping* 34.9, pp. 2154–2177. ISSN: 10659471. DOI: 10.1002/hbm.22058.

Jackson, Iain and Sylvain Sirois (2009). "Infant cognition: Going full factorial with pupil dilation". In: *Developmental Science* 12.4, pp. 670–679. ISSN: 1363755X. DOI: 10.1111/j.1467-7687.2008.00805.x.

Joanisse, Marc F. and James L. McClelland (2015). "Connectionist perspectives on language learning, representation and processing". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 6.3, pp. 235–247. ISSN: 19395086. DOI: 10.1002/wcs.1340.

Johnson, Elizabeth K. and Peter W. Jusczyk (2001). "Word Segmentation by 8-Month-Olds: When Speech Cues Count More Than Statistics". In: *Journal of*

Memory and Language 44, pp. 548–567. ISSN: 0749596X. DOI: 10.1006/jmla.2000.2755.

Johnson, Elizabeth K. and Michael D. Tyler (2010). "Testing the limits of statistical learning for word segmentation". In: *Developmental Science* 13.2, pp. 339–345. ISSN: 1363755X. DOI: 10.1111/j.1467-7687.2009.00886.x.

Johnson, Mark H. (2001). "Functional Brain Development in Humans". In: *Nature Reviews Neuroscience* 2.7, pp. 475–483. ISSN: 1471003X. DOI: 10.1038/35081509.

Johnson, Scott P et al. (2009). "Abstract rule learning for visual sequences in 8- and 11-month-olds". In: *Infancy : the official journal of the International Society on Infant Studies* 14.1, pp. 2–18. ISSN: 1525-0008. DOI: 10.1080/15250000802569611.

Jones, David T. et al. (2012). "Non-stationarity in the "resting brain's" modular architecture". In: *PLoS ONE* 7.6. ISSN: 19326203. DOI: 10.1371/journal.pone.0039731.

Jusczyl, Peter W., David B. Pisoni, and John Mullennix (1992). "Some consequences of stimulus variability on speech processing by 2-month-old infants". In: *Cognition* 43.3, pp. 253–291. ISSN: 00100277. DOI: 10.1016/0010-0277(92)90014-9.

Kabdebon, C. et al. (2015). "Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants". In: *Brain and Language* 148, pp. 25–36. ISSN: 10902155. DOI: 10.1016/j.bandl.2015.03.005.

Kahneman, Daniel and Jackson Beatty (1966). "Pupil Diameter and Load on Memory". In: *Science* 154.3756, pp. 1583–1585. ISSN: 0036-8075. DOI: 10.1126/science.154.3756.1583.

Karuza, Elisabeth A. et al. (2013). "The neural correlates of statistical learning in a word segmentation task: An fMRI study". In: *Brain and Language* 127.1, pp. 46–54. ISSN: 0093934X. DOI: 10.1016/j.bandl.2012.11.007. arXiv: NIHMS150003.

Kirkham, N Z et al. (2002). "Visual statistical learning in infancy: Evidence for a domain-general learning mechanism". In: *Cognition* 83, pp. 4–5.

Kisilevsky, B. S. et al. (2004). "Maturation of fetal responses to music". In: *Developmental Science* 7.5, pp. 550–559. ISSN: 1363755X. DOI: 10.1111/j.1467-7687.2004.00379.x.

Kudo, Noriko et al. (2011). "On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials". In: *Developmental Science* 14.5, pp. 1100–1106. ISSN: 1363755X. DOI: 10.1111/j.1467-7687.2011.01056.x.

Kuhl, P. et al. (1992). "Linguistic experience alters phonetic perception in infants by 6 months of age". In: *Science* 255.5044, pp. 606–608. ISSN: 0036-8075. DOI: 10.1126/science.1736364.

Kuhl, Patricia K. (1983). "Perception of auditory equivalence classes for speech in early infancy". In: *Infant Behavior and Development* 6.2-3, pp. 263–285. ISSN: 01636383. DOI: 10.1016/S0163-6383(83)80036-8.

Laeng, B., S. Sirois, and G. Gredeback (2012). "Pupillometry: A Window to the Preconscious?" In: *Perspectives on Psychological Science* 7.1, pp. 18–27. ISSN: 1745-6916. DOI: 10.1177/1745691611427305.

Langus, Alan et al. (2012). "Can prosody be used to discover hierarchical structure in continuous speech?" In: *Journal of Memory and Language* 66.1, pp. 285–306. ISSN: 0749596X. DOI: 10.1016/j.jml.2011.09.004.

Larson-Prior, Linda J. et al. (2011). "Modulation of the brain's functional network architecture in the transition from wake to sleep". In: *Progress in Brain Research* 193, pp. 277–294. ISSN: 00796123. DOI: 10.1016/B978-0-444-53839-0.00018-1.

Leonardi, Nora et al. (2013). "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest". In: *NeuroImage* 83, pp. 937–950. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.07.019.

Leroy, François et al. (2011). ""Early Maduration of the Linguistic Dorsal Pathway in Human Infants"". In: *The Journal of Neuroscience* 31.4, pp. 1500–1506. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4141-10.2011.

Lewkowicz, David J and Amy M Hansen-Tift (2012). "Infants deploy selective attention to the mouth of a talking face when learning speech." In: *Proceedings of the National Academy of Sciences of the United States of America* 109.5, pp. 1431–6. ISSN: 1091-6490. DOI: 10.1073/pnas.1114783109.

Lichtman, J. W. and W. Denk (2011). "The Big and the Small: Challenges of Imaging the Brain's Circuits". In: *Science* 334.6056, pp. 618–623. ISSN: 0036-8075. DOI: 10.1126/science.1209168.

Lieberman, P (1984). *The biology and evolution of language*, p. 379. ISBN: 0674074122.

Liu, Yan et al. (2010). "Structural asymmetries in motor and language networks in a population of healthy preterm neonates at term equivalent age: A diffusion tensor imaging and probabilistic tractography study". In: *NeuroImage* 51.2, pp. 783–788. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.02.066.

Lloyd-Fox, S., A. Blasi, and C.E. E. Elwell (2010). "Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy". In: *Neuroscience and Biobehavioral Reviews* 34.3, pp. 269–284. ISSN: 01497634. DOI: 10.1016/j.neubiorev.2009.07.008.

Lu, Chun Ming et al. (2010). "Use of fNIRS to assess resting state functional connectivity". In: *Journal of Neuroscience Methods* 186.2, pp. 242–249. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2009.11.010.

Lynall, Mary-Ellen et al. (2010). "Functional connectivity and brain networks in schizophrenia." In: *The Journal of neuroscience : the official journal of the Society for*

Neuroscience 30.28, pp. 9477–87. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0333-10.2010.

Mahmoudzadeh, Mahdi et al. (2013). "Syllabic discrimination in premature human infants prior to complete formation of cortical layers." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.12, pp. 4846–51. ISSN: 1091-6490. DOI: 10.1073/pnas.1212220110.

Majeed, Waqas et al. (2011). "Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans". In: *NeuroImage* 54.2, pp. 1140–1150. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.08.030.

Maye, Jessica, Janet F. Werker, and LouAnn Gerken (2002). "Infant sensitivity to distributional information can affect phonetic discrimination". In: *Cognition* 82.3, pp. 101–111. ISSN: 00100277. DOI: 10.1016/S0010-0277(01)00157-3.

McAvoy, M. et al. (2012). "Dissociated mean and functional connectivity BOLD signals in visual cortex during eyes closed and fixation". In: *Journal of Neurophysiology* 108.9, pp. 2363–2372. ISSN: 0022-3077. DOI: 10.1152/jn.00900.2011.

McNealy, K., J. C. Mazziotta, and M. Dapretto (2006). "Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing". In: *Journal of Neuroscience* 26.29, pp. 7629–7639. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5501-05.2006.

Medvedev, Andrei V. (2014). "Does the resting state connectivity have hemispheric asymmetry? A near-infrared spectroscopy study". In: *NeuroImage* 85, pp. 400–407. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.05.092. arXiv: NIHMS150003.

Mehler, Jacques et al. (1988). "A precursor of language acquisition in young infants". In: *Cognition* 29.2, pp. 143–178. ISSN: 00100277. DOI: 10.1016/0010-0277(88)90035-2.

Meyer, Martin et al. (2002). *FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences*. DOI: 10.1002/hbm.10042.

Minagawa-Kawai, Yasuyo et al. (2008). "Optical imaging of infants' neurocognitive development: Recent advances and perspectives". In: *Developmental Neurobiology* 68.6, pp. 712–728. ISSN: 19328451. DOI: 10.1002/dneu.20618.

Minagawa-Kawai, Yasuyo et al. (2011). "Optical brain imaging reveals general auditory and language-specific processing in early infant development". In: *Cerebral Cortex* 21.2, pp. 254–261. ISSN: 10473211. DOI: 10.1093/cercor/bhq082.

Molavi, Behnam et al. (2013). "Analyzing the resting state functional connectivity in the human language system using near infrared spectroscopy." In: *Frontiers in human neuroscience* 7.January, p. 921. ISSN: 1662-5161. DOI: 10.3389/fnhum.2013.00921.

Murdock, Bennet B (1962). "The serial position effect of free recall." In: *Journal of Experimental Psychology* 64.5, pp. 482–488. ISSN: 00221015. DOI: 10.1037/h0045106.

Musso, Mariacristina et al. (2003). "Broca's area and the language instinct." In: *Nature neuroscience* 6.7, pp. 774–781. ISSN: 10976256. DOI: 10.1038/nn1077.

Nakano, Tamami et al. (2009). "Prefrontal cortical involvement in young infants' analysis of novelty." In: *Cerebral cortex (New York, N.Y. : 1991)* 19.2, pp. 455–63. ISSN: 1460-2199. DOI: 10.1093/cercor/bhn096.

Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler (1998). "Language discrimination by newborns: Toward an understanding of the role of rhythm." In: *Journal of Experimental Psychology: Human Perception and Performance* 24.3, pp. 756–766. ISSN: 1939-1277. DOI: 10.1037/0096-1523.24.3.756.

Nazzi, Thierry, Peter W Jusczyk, and Elizabeth K Johnson (2000). "Language Discrimination by English-Learning 5-Month-Olds: Effects of Rhythm and Familiarity". In: *Journal of Memory and Language* 43.1, pp. 1–19. ISSN: 0749596X. DOI: 10.1006/jmla.2000.2698.

Nelson, Deborah G. Kemler et al. (1989). "How the prosodic cues in motherese might assist language learning". In: *Journal of Child Language* 16.01, p. 55. ISSN: 0305-0009. DOI: 10.1017/S030500090001343X.

Nespor, Marina and Irene Vogel (2007). *Prosodic phonology :with a new foreword*.

Nespor, Marina et al. (2008). "Different Phrasal Prominence Realizations in VO and OV Languages". In: *Lingue e Linguaggio* VII.2, pp. 1–29. ISSN: 17209331. DOI: 10.1418/28093.

Newport, Elissa L (2002). "Critical periods in language development". In: *Encyclopedia of Cognitive Science*, pp. 737–740.

Niu, H and Y He (2014). "Resting-state functional brain connectivity: lessons from functional near-infrared spectroscopy". In: *Neuroscientist* 20.2, pp. 173–188. ISSN: 1073-8584. DOI: 10.1177/1073858413502707.

Niu, Haijing et al. (2012). "Revealing Topological Organization of Human Brain Functional Networks with Resting-State Functional near Infrared Spectroscopy". In: *PLoS ONE* 7.9, pp. 1–14. ISSN: 19326203. DOI: 10.1371/journal.pone.0045771.

Oostenveld, Robert et al. (2011). "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data". In: *Computational Intelligence and Neuroscience* 2011, pp. 1–9. ISSN: 1687-5265. DOI: 10.1155/2011/156869. arXiv: 156869.

Park, Bumhee et al. (2012). "Are brain networks stable during a 24-hour period?" In: *NeuroImage* 59.1, pp. 456–466. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2011.07.049.

Park, H.-J. and Karl Friston (2013). "Structural and Functional Brain Networks: From Connections to Cognition". In: *Science* 342.6158, pp. 1238411–1238411. ISSN: 0036-8075. DOI: 10.1126/science.1238411.

Peña, Marcela, Ricardo a H Bion, and Marina Nespor (2011). "How modality specific is the iambic-trochaic law? Evidence from vision." In: *Journal of experimental psychology. Learning, memory, and cognition* 37.5, pp. 1199–208. ISSN: 1939-1285. DOI: 10.1037/a0023944.

Peña, Marcela et al. (2002). "Signal-driven computations in speech processing." In: *Science (New York, N.Y.)* 298.5593, pp. 604–7. ISSN: 1095-9203. DOI: 10.1126/science.1072901.

Peña, Marcela et al. (2003). "Sounds and silence: an optical topography study of language recognition at birth." In: *Proceedings of the National Academy of Sciences of the United States of America* 100.OCTOBER, pp. 11702–11705. ISSN: 0027-8424. DOI: 10.1073/pnas.1934290100.

Perani, Daniela et al. (2011). "Neural language networks at birth". In: *Proceedings of the National Academy of Sciences* 108.38, pp. 1–6. ISSN: 0027-8424. DOI: 10.1073/pnas.1102991108.

Perruchet, Pierre and Sebastien Pacton (2006). "Implicit learning and statistical learning: one phenomenon, two approaches". In: *Trends in Cognitive Sciences* 10.5, pp. 233–238. ISSN: 13646613. DOI: 10.1016/j.tics.2006.03.006.

Perruchet, Pierre and Annie Vinter (1998). "PARSER: A Model for Word Segmentation". In: *Journal of Memory and Language* 39, pp. 246–263. ISSN: 0749596X. DOI: 10.1006/jmla.1998.2576.

Petersen, Steven E. and Olaf Sporns (2015). "Brain Networks and Cognitive Architectures". In: *Neuron* 88.1, pp. 207–219. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.09.027.

Poeppel, David (2003). "The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'". In: *Speech Communication*. Vol. 41. 1, pp. 245–255. ISBN: 0167-6393. DOI: 10.1016/S0167-6393(02)00107-3.

Polka, Linda and Janet F Werker (1994). "Developmental changes in perception of nonnative vowel contrasts." In: *Journal of Experimental Psychology: Human Perception and Performance* 20.2, pp. 421–435. ISSN: 1939-1277. DOI: 10.1037/0096-1523.20.2.421.

Power, Jonathan D., Bradley L. Schlaggar, and Steven E. Petersen (2014). *Studying brain organization via spontaneous fMRI signal*. DOI: 10.1016/j.neuron.2014.09.007. arXiv: NIHMS150003.

Power, Jonathan D. et al. (2011). "Functional Network Organization of the Human Brain". In: *Neuron* 72.4, pp. 665–678. ISSN: 08966273. DOI: 10.1016/j.neuron.2011.09.006.

Power, Jonathan D. et al. (2012). "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion". In: *NeuroImage* 59.3, pp. 2142–2154. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2011.10.018. arXiv: NIHMS150003.

Price, Cathy J. (2010). "The anatomy of language: A review of 100 fMRI studies published in 2009". In: *Annals of the New York Academy of Sciences* 1191, pp. 62–88. ISSN: 00778923. DOI: 10.1111/j.1749-6632.2010.05444.x.

Querleu, Denis et al. (1984). "Reaction of the newborn infant less than 2 hours after birth to the maternal voice". In: *Journal de Gynécologie Obstetrique et Biologie de la Réproduction* 13.2, pp. 125–134. ISSN: 0368-2315.

Rubinov, Mikail and Olaf Sporns (2010). "Complex network measures of brain connectivity: Uses and interpretations". In: *NeuroImage* 52.3, pp. 1059–1069. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2009.10.003.

Ruiter, J P de and Stephen C Levinson (2008). "A Biological Infrastructure for Communication Underlies the Cultural Evolution of Languages". In: *Behavioral and Brain Sciences* 31.5, p. 518. ISSN: 0140-525X. DOI: doi:10.1017/S0140525X08005086.

Saez de Urabain, Irati R, Mark H Johnson, and Tim J Smith (2015). "GraFIX: A semiautomatic approach for parsing low- and high-quality eye-tracking data". In: *Behavior Research Methods* 47.1, pp. 53–72. ISSN: 1554-3528. DOI: 10.3758/s13428-014-0456-0.

Saffran, Jenny R., Richard N. Aslin, and Elissa L. Newport (1996). *Statistical learning by 8-month-old infants*. DOI: 10.1126/science.274.5294.1926.

Saffran, Jenny R., Elissa L. Newport, and Richard N. Aslin (1996). "Word Segmentation: The Role of Distributional Cues". In: *Journal of Memory and Language* 35.4, pp. 606–621. ISSN: 0749596X. DOI: 10.1006/jmla.1996.0032.

Saffran, Jenny R. et al. (1999). "Statistical learning of tone sequences by human infants and adults". In: *Cognition* 70.1, pp. 27–52. ISSN: 00100277. DOI: 10.1016/S0010-0277(98)00075-4.

Saffran, Jenny R. et al. (2007). "Dog is a dog is a dog: Infant rule learning is not specific to language". In: *Cognition* 105.3, pp. 669–680. ISSN: 00100277. DOI: 10.1016/j.cognition.2006.11.004.

Sasai, Shuntaro et al. (2011). "Frequency-specific functional connectivity in the brain during resting state revealed by NIRS". In: *NeuroImage* 56.1, pp. 252–257. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.12.075.

Sato, Yutaka et al. (2013). "Dialectal differences in hemispheric specialization for Japanese lexical pitch accent". In: *Brain and Language* 127.3, pp. 475–483. ISSN: 0093934X. DOI: 10.1016/j.bandl.2013.09.008.

Schapiro, Anna C et al. (2014). "The Necessity of the Medial Temporal Lobe for Statistical Learning". In: *Journal of Cognitive Neuroscience* 26.8, pp. 1736–1747. ISSN: 0898-929X. DOI: 10.1162/jocn_a_00578.

Scholkmann, F et al. (2010). "How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation." In: *Physiological measurement* 31.5, pp. 649–62. ISSN: 1361-6579. DOI: 10.1088/0967-3334/31/5/004.

Scholkmann, Felix et al. (2014). *A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology*. DOI: 10.1016/j.neuroimage.2013.05.004.

Seidl, Amanda and Elizabeth K. Johnson (2006). "Infant word segmentation revisited: Edge alignment facilitates target extraction". In: *Developmental Science* 9.6, pp. 565–573. ISSN: 1363755X. DOI: 10.1111/j.1467-7687.2006.00534.x.

Shukla, Mohinish, Marina Nespor, and Jacques Mehler (2007). "An interaction between prosody and statistics in the segmentation of fluent speech." In: *Cognitive psychology* 54.1, pp. 1–32. ISSN: 0010-0285. DOI: 10.1016/j.cogpsych.2006.04.002.

Shukla, Mohinish, Katherine S White, and Richard N Aslin (2011). "Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.15, pp. 6038–43. ISSN: 1091-6490. DOI: 10.1073/pnas.1017617108.

Sigurd, B, M Eeg-Olofsson, and J van de Weijer (2004). "Word length, sentence length and frequency - Zipf revisited". In: *Studia Linguistica* 58.1, pp. 37–52. ISSN: 0039-3193. DOI: 10.1111/j.0039-3193.2004.00109.x.

Singh, Archana K. and Ippeita Dan (2006). "Exploring the false discovery rate in multichannel NIRS". In: *NeuroImage* 33.2, pp. 542–549. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2006.06.047.

Sirois, Sylvain and Iain R. Jackson (2012). "Pupil dilation and object permanence in infants". In: *Infancy* 17.1, pp. 61–78. ISSN: 15250008. DOI: 10.1111/j.1532-7078.2011.00096.x.

Skinner, B F (2003). "Verbal behavior". In: *The history of psychology: Fundamental questions*, pp. 399–407. ISBN: 0-19-515154-2 (Paperback). DOI: 10.1037/11256-000.

Smith, Stephen M et al. (2009). "Correspondence of the brain's functional architecture during activation and rest." In: *Proceedings of the National Academy of Sciences of the United States of America* 106.31, pp. 13040–5. ISSN: 1091-6490. DOI: 10.1073/pnas.0905267106. arXiv: 0905267106.

Soderstrom, Melanie et al. (2003). *The prosodic bootstrapping of phrases: Evidence from prelinguistic infants*. DOI: 10.1016/S0749-596X(03)00024-X.

Stiles, Joan and Terry L. Jernigan (2010). "The basics of brain development". In: *Neuropsychology Review* 20.4, pp. 327–348. ISSN: 10407308. DOI: 10.1007/s11065-010-9148-4.

Strangman, Gary, Maria Angela Franceschini, and David A. Boas (2003). "Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters". In: *NeuroImage* 18.4, pp. 865–879. ISSN: 10538119. DOI: 10.1016/S1053-8119(03)00021-1.

Streeter, L a (1976). "Language perception of 2-month-old infants shows effects of both innate mechanisms and experience." In: *Nature* 259.5538, pp. 39–41. ISSN: 0028-0836. DOI: 10.1038/259039a0.

Swingley, Daniel (2005). "Statistical clustering and the contents of the infant vocabulary". In: *Cognitive Psychology* 50.1, pp. 86–132. ISSN: 00100285. DOI: 10.1016/j.cogpsych.2004.06.001.

Teinonen, Tuomas et al. (2009). "Statistical language learning in neonates revealed by event-related brain potentials". In: *BMC Neuroscience* 10, p. 21. ISSN: 1471-2202. DOI: 10.1186/1471-2202-10-21.

Telkemeyer, Silke et al. (2009). "Sensitivity of newborn auditory cortex to the temporal structure of sounds." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29.47, pp. 14726–33. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.1246-09.2009.

Thiessen, Erik D., Emily A. Hill, and Jenny R. Saffran (2005). "Infant-directed speech facilitates word segmentation". In: *Infancy* 7.1, pp. 53–71. ISSN: 15250008. DOI: 10.1207/s15327078in0701_5.

Thiessen, Erik D. and Jenny R. Saffran (2003). "When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants." In: *Developmental Psychology* 39.4, pp. 706–716. ISSN: 0012-1649. DOI: 10.1037/0012-1649.39.4.706.

— (2007). *Learning to Learn: Infants' Acquisition of Stress- Based Strategies for Word Segmentation*.

Toro, Juan M., Scott Sinnett, and Salvador Soto-Faraco (2005). "Speech segmentation by statistical learning depends on attention". In: *Cognition* 97.2, pp. 25–34. ISSN: 00100277. DOI: 10.1016/j.cognition.2005.01.006.

Toro, Juan M and Josep B Trobalón (2005). "Statistical computations over a speech stream in a rodent." In: *Perception & psychophysics* 67.5, pp. 867–875. ISSN: 0031-5117. DOI: 10.3758/BF03193539.

Turk-Browne, Nicholas B. et al. (2009). "Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness". In: *Journal of Cognitive Neuroscience* 21.10, pp. 1934–1945. ISSN: 0898-929X. DOI: 10.1162/jocn.2009.21131.

Wartenburger, Isabell et al. (2007). "The processing of prosody: Evidence of interhemispheric specialization at the age of four." In: *NeuroImage* 34.1, pp. 416–25. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.09.009.

Wass, S V, T J Smith, and M H Johnson (2013). "Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults". In: *Behavior Research Methods* 45.1, pp. 229–250. ISSN: 1554-3528. DOI: 10.3758/s13428-012-0245-6.

Werker, Janet F. and Richard C. Tees (2002). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life". In: *Infant*

Behavior and Development 25.1, pp. 121–133. ISSN: 01636383. DOI: 10.1016/S0163-6383(02)00093-0.

Werker, Janet F. et al. (1981). "Developmental aspects of cross-language speech perception." In: *Child development* 52.1, pp. 349–355. ISSN: 0009-3920. DOI: 10.2307/1129249.

White, Brian R. et al. (2012). "Bedside optical imaging of occipital resting-state functional connectivity in neonates". In: *NeuroImage* 59.3, pp. 2529–2538. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2011.08.094.

Wijk, Bernadette C M van, Cornelis J. Stam, and Andreas Daffertshofer (2010). "Comparing brain networks of different size and connectivity density using graph theory". In: *PLoS ONE* 5.10. ISSN: 19326203. DOI: 10.1371/journal.pone.0013701.

Yan, Chao Gan et al. (2013). "A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics". In: *NeuroImage* 76, pp. 183–201. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2013.03.004. arXiv: NIHMS150003.

Yang, Charles D. (2004). *Universal Grammar, statistics or both?* DOI: 10.1016/j.tics.2004.08.006.

Yang, Jiale et al. (2016). "Cortical response to categorical color perception in infants investigated by near-infrared spectroscopy". In: *Proceedings of the National Academy of Sciences* 113.9, pp. 2370–2375. ISSN: 0027-8424. DOI: 10.1073/pnas.1512044113.

Yoshida, Katherine A. et al. (2010). "The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study". In: *Cognition* 115.2, pp. 356–361. ISSN: 00100277. DOI: 10.1016/j.cognition.2010.01.005.

Yucel, M A et al. (2014). "Targeted principle component analysis: A new motion artifact correction approach for near-infrared spectroscopy". In: *Journal of Innovative Optical Health Sciences* 7.2, pp. 1–8. ISSN: 1793-5458. DOI: 10.1142/S1793545813500661.

Zalesky, Andrew et al. (2014). "Time-resolved resting-state brain networks." In: *Proceedings of the National Academy of Sciences of the United States of America* 111.28, pp. 10341–6. ISSN: 1091-6490. DOI: 10.1073/pnas.1400181111.

Zatorre, R J and P Belin (2001). "Spectral and temporal processing in human auditory cortex." In: *Cerebral cortex (New York, N.Y. : 1991)* 11.10, pp. 946–53. ISSN: 1047-3211. DOI: 10.1093/cercor/11.10.946.

Zhang, Han et al. (2010). "Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements". In: *NeuroImage* 51.3, pp. 1150–1161. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2010.02.080.

Zhang, Han et al. (2011). "Test-retest assessment of independent component analysis-derived resting-state functional connectivity based on functional near-infrared

spectroscopy." In: *NeuroImage* 55.2, pp. 607–615. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2010.12.007.

Zhang, Yu-Jin et al. (2014). "Detecting resting-state functional connectivity in the language system using functional near-infrared spectroscopy." In: *Journal of biomedical optics* 15.4, p. 047003. ISSN: 1560-2281. DOI: 10.1117/1.3462973.